(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 January 2001 (04.01.2001)

PCT

(10) International Publication Number
WO 01/00849 A1

(51) International Patent Classification[7]: C12N 15/54, 9/10, C07K 16/40, C12Q 1/68, 1/48, A61K 38/45 // A61P 35/00

(21) International Application Number: PCT/US00/17827

(22) International Filing Date: 28 June 2000 (28.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/141,582 29 June 1999 (29.06.1999) US

(71) Applicant: ICOS CORPORATION [US/US]; 22021 20th Avenue S.E., Bothell, WA 98021 (US).

(72) Inventors: CHRISTENSON, Erik; 10610 S.E. 25th Street, Bellevue, WA 98004 (US). DEMAGGIO, Anthony, J.; 12904 126th Court N.E., Kirkland, WA 98034 (US). GOLDMAN, Phyllis, S.; 3903 243rd Place S.E. N-301, Bothell, WA 98021 (US). MCELLIGOTT, David, L.; 19621 97th Avenue N.E., Bothell, WA 98011 (US).

(74) Agent: NOLAND, Greta, E.; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606-6402 (US).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/00849 A1

(54) Title: TANKYRASE2 MATERIALS AND METHODS

(57) Abstract: The invention provides novel tankyrase polypeptides designated tankyrase2, polynucleotides encoding the polypeptides, expression constructs comprising the polynucleotides, and host cells transformed with the expression constructs. Also provided are methods for producing the tankyrase2 polypeptides, antibodies that are immunoreactive with the tankyrase2 polypeptides. In addition, there are provided methods for identifying specific binding partners of tankyrase2, and more particularly methods for identifying binding partners that modulate biological activity of tankyrase2. Methods of modulating biological activity of tankyrase2 *in vitro* and *in vivo* are also provided.

TANKYRASE2 MATERIALS AND METHODS

This application claims the benefit of United States Provisional Application Serial No. 60/141,582, filed June 29, 1999.

The present invention relates generally to a novel tankyrase polypeptide having poly ADP-ribosylation activity, to polynucleotides encoding the polypeptide, and to methods of using such materials.

BACKGROUND OF THE INVENTION

The ends of eukaryotic chromosomes (telomeres) are characterized by simple repeat DNA sequences. The length and sequence of the repeats varies from species to species but the importance of telomeres is universal in organisms with linear chromosomes. Telomeres protect the ends of the chromosomes and ostensibly function to prevent recombination of chromosome ends, which leads to chromosomal fusion and instability. In addition, there is considerable evidence that the length of the telomere repeats determines the ability of a cell to divide or perhaps even to survive.

The telomeres of cultured primary human fibroblasts become progressively shorter with each cell division in the absence of an active mechanism to regenerate telomere length [Harley et al., *Nature* 345:458-60 (1990)]. At some critical stage of telomere shortening, these cells are no longer able to divide and enter a state known as cellular senescence. Thus, in human primary fibroblasts at least, telomere length functions as a biological clock to monitor cellular aging and regulate longevity.

The observation that telomere length regulates cellular aging prompted the hypothesis that telomere regulation may also be critical for organismal aging. Mice that are unable to replicate telomeres show characteristics of premature aging after the third generation. These characteristics include premature graying, decreased cell division capacity, impaired wound healing, and increased cancer incidence amongst others. Thus, regulation of telomere structure may be critical for some of the characteristics associated with aging. Drugs that modulate the regulation of telomere structure thus may have utility in treatment of age-related syndromes or in cases of genetically determined premature aging syndromes.

Only recently has some of the machinery that replicates telomeres been described. This machinery, collectively referred to as the telomerase complex, consists of several proteins that replicate the telomeres and protect the telomere structure from DNA repair, which otherwise might treat telomeres as damaged DNA and affect end joining or recombination thus destroying the integrity of the chromosome. Telomerase is the replicative component of the telomerase complex and is a DNA polymerase that features an integral RNA molecule that serves as the template for the addition of the repetitive sequences [for a review, see Greider, *Ann Rev Biochem* 65:337-65 (1996)]. The observation that telomerase activity is essential for continued cell division suggests that inappropriate telomerase activity may be, in some instances, a contributing factor in the oncogenic transformation of cells. Forced expression of telomerase does not in and of itself cause oncogenic transformation but the fact that cells overexpressing telomerase have apparently unlimited capacity to replicate suggests that inappropriate expression of telomerase may be one step in a multi-step process of oncogenic transformation. In addition, numerous studies have shown that telomerase activity is higher in tumor tissue than most normal tissues suggesting that increased telomerase activity may be essential for tumor growth [for reviews, see Bacchetti, *Cancer Surv* 28:197-216 (1996); and Harley et al., *Cold Spring Harbor Symp Quant Biol* 59:307-15(1994)].

Two telomere-specific DNA binding proteins, designated TRF1 and TRF2 have also been shown to be important for maintenance of telomeres [Chong et al., *Science* 270:1663-7 (1995); van Steensel et al., *Cell* 92:401-13 (1998)]. TRF1 has a critical role in the regulation of telomere length while TRF2 seems to be important for protecting chromosome ends. Both molecules contain DNA binding domains and dimerization domains and both appear to function as homodimers. Binding of TRF1 to telomere repeats inhibits the function of telomerase thus contributing to telomere shortening during replication [van Steensel and de Lange, *Nature* 385:740-3 (1997)].

An additional molecule, tankyrase, has been identified which modifies TRF1 by the addition of polymers of ADP-ribose [Smith et al., *Science* 282:1484-7 (1998)]. Tankyrase is structurally and functionally related to the Poly(ADP-Ribose) Polymerase (PARP) molecule, which modifies proteins by the addition of ADP-ribose

polymers [for review see Alvarez-Gonzalez et al., *Mol Cell Biochem* 138:33-7 (1994)]. The structural relationship to PARP exists in a putative catalytic domain of tankyrase that has extensive amino acid sequence similarity to PARP. In addition, tankyrase contains a sterile alpha motif (SAM) and 24 ankyrin repeats. These structures are typically involved in protein/protein interactions and at least a portion of the ankyrin repeat region in tankyrase has been shown to be responsible for the interaction with TRF1. Tankyrase has been shown to poly ADP-ribosylate TRF1 *in vitro* and it has been suggested that the role of tankyrase *in vivo* is to ADP-ribosylate TRF1 causing dissociation of TRF1 from the telomere repeats and thus allowing telomerase to replicate the telomeres. Drugs that inhibit tankyrase activity then might be expected to inhibit the replication of telomeres and thus cause eventual senescence of dividing cell populations such as cancer cells or proliferating immune system cell as examples.

As tankyrase or tankyrase-related gene products might be attractive targets of drug design, there is a need in the art to identify additional molecules with related functions and/or structures. Such molecules might serve as specificity controls for tankyrase targeted drugs or may themselves be suitable targets for drug discovery programs.

In view of the above considerations, it is clear that existing knowledge is lacking with respect to cellular DNA repair mechanisms, signaling, and induction of cellular replication, mechanisms of tumorigenesis, and treatment of cancer disease states. Thus, there exists a need in the art for the identification of additional tankyrase-like molecules for use in determining the selectivity of therapeutics designed to modulate tankyrase function and as targets in their own right for therapeutic intervention in human diseases. The profiling of tankyrase inhibitors on additional tankyrase gene products may allow for the tankyrase-selective drugs, which could be beneficial for particular indications, the reduction of undesirable side effects, or the targeting of therapeutics to selected tissues. Other purposes and advantages of the invention will be readily apparent to the artisan having ordinary skill in the art.

SUMMARY OF THE INVENTION

It has now been discovered that these and other purposes can be achieved by the present invention, which, in one aspect, provides purified and isolated tankyrase2 polypeptides, preferably human tankyrase2 polypeptides. In particular the invention provides a purified and isolated tankyrase2 polypeptide comprising the amino acid sequence defined in SEQ ID NO:133 (designated "TANK2-LONG") or SEQ ID NO:135 (designated "TANK2-SHORT"). The invention also provides polynucleotides encoding the tankyrase2 polypeptides. For example, the polynucleotide may comprise the coding region of the nucleotide sequence defined in SEQ ID NO:132 or SEQ ID NO:134.

The invention further provides polynucleotides that are complements to TANK2-encoding polynucleotides, as well as polynucleotides that hybridize under moderately stringent hybridization conditions to the coding or non-coding strand of the tankyrase2 polynucleotides. In a preferred case, the polynucleotide hybridizes to the complement of the polynucleotide defined in SEQ ID NO:132 or SEQ ID NO:134 under stringent hybridization conditions, and encodes a protein that: (a) has poly(ADP) polymerase activity, (b) interacts with damaged DNA, or (c) binds to telomere repeat-binding factors and/or modulates their activity.

The polynucleotides may be DNA molecules or RNA molecules. Certain desirable polynucleotides of the invention, e.g., oligonucleotide probes, may further comprise a detectable label moiety.

In another aspect, the invention provides an expression construct, comprising a tankyrase2-encoding polynucleotide, as well as host cells transformed or transfected with the expression constructs. The polynucleotide can be operatively linked to a heterologous promoter.

In a further aspect, the invention provides a method for producing a tankyrase2 polypeptide in a host cell modified to express the tankyrase polypeptide, comprising the steps of:

a) growing the host cell under conditions appropriate for expression of the tankyrase2 polypeptide; and

b) isolating the tankyrase2 polypeptide from the host cell or the medium in which the host cell is grown.

In yet another aspect, the invention provides antibodies that are immunoreactive with a tankyrase2 polypeptide. For example, the antibodies may be selected from the group consisting of monoclonal antibodies, polyclonal antibodies, single chain antibodies (scFv antibodies), chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, CDR-grafted antibodies, Fab fragments, Fab' fragments, $F(ab')_2$ fragments, and Fv fragments. Also provided are cell lines that produce such antibodies. There are also provided anti-idiotype antibodies that are immunoreactive with tankyrase2-specific antibodies.

In still another aspect, the invention provides a method for identifying a binding partner of a tankyrase2 polypeptide, comprising:

a) contacting the tankyrase2 polypeptide with a test compound under conditions that permit binding of the tankyrase2 polypeptide and the test compound;

b) detecting binding of the test compound and the tankyrase2 polypeptide; and

c) identifying the test compound as a binding partner of the tankyrase2 polypeptide.

For example, the method can be used to identify binding partners that selectively or specifically modulate, i.e., inhibit or enhance, a biological activity of the tankyrase2 polypeptide.

Also provided in another aspect is a method for identifying a binding partner of a tankyrase2 polynucleotide, comprising:

a) contacting the tankyrase2 polynucleotide with a test compound under conditions that permit binding of the tankyrase2 polynucleotide and the test compound;

b) detecting binding of the test compound and the tankyrase2 polynucleotide; and

c) identifying the test compound as a binding partner of the tankyrase2 polynucleotide.

The method may be used to identify binding partners that selectively or specifically modulate, i.e., inhibit or enhance, expression of the tankyrase2 polypeptide.

There is also provided by the invention a method of treating a human or animal subject having a medical condition mediated by poly(ADP-ribose) polymerase activity, comprising administering to the subject a tankyrase2 inhibitory compound in an amount effective for inhibiting tankyrase2 in the subject. In another aspect, the invention provides a method of treating a human or animal subject having a medical condition mediated by poly(ADP-ribose) polymerase activity, comprising administering to the subject a compound that inhibits tankyrase2 expression or activity in an amount effective for inhibiting poly(ADP-ribose) polymerase activity in the subject. The method is of particular interest in treating medical conditions associated with growth of neoplastic tissue. For example, the method can be used to treat cancers such as carcinomas, sarcomas, leukemias, and lymphomas. More particularly, the method may be used to treat cancers selected from the group consisting of ACTH-producing tumor, acute lymphocytic leukemia, acute nonlymphocytic leukemia, cancer of the adrenal cortex, bladder cancer, brain cancer, breast cancer, cervical cancer, chronic lymphocytic leukemia, chronic myelocytic leukemia, colorectal cancer, cutaneous T-cell lymphoma, endometrial cancer, esophageal cancer, Ewing's sarcoma, gallbladder cancer, hairy cell leukemia, head and neck cancer, Hodgkin's lymphoma, Kaposi's sarcoma, kidney cancer, liver cancer, lung cancer (small and non-small cell), malignant peritoneal effusion, malignant pleural effusion, melanoma, mesothelioma, multiple myeloma, neuroblastoma, glioma, non-Hodgkin's lymphoma, osteosarcoma, ovarian cancer, ovarian (germ cell) cancer, pancreatic cancer, penile cancer, prostate cancer, retinoblastoma, skin cancer, soft tissue sarcoma, squamous cell carcinomas, stomach cancer, testicular cancer, thyroid cancer, trophoblastic neoplasms, uterine cancer, vaginal cancer, cancer of the vulva, and Wilm's tumor.

These and other features and advantages of the present invention will be appreciated from the detailed description and examples that are set forth herein. The detailed description and examples are provided to enhance the understanding of the invention, but are not intended to limit the scope of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates generally to a previously uncharacterized nucleic acid encoding a novel human protein designated "tankyrase2" (hereinafter also referred to as "TANK2"). As illustrated herein tankyrase2 is distinct from known tankyrase proteins and other proteins sharing poly(ADP-ribose) polymerase activity. The present invention is based on the discovery of novel gene encoding the tankyrase2 protein, and nucleic acid sequences, oligonucleotides, fragments, and antisense molecules thereof.

The nucleotide sequence information provided by the invention makes possible large-scale expression of the encoded TANK2 polypeptide by techniques well known and routinely practiced in the art. The invention also permits identification and isolation of polynucleotides encoding related TANK2 polypeptides by well-known techniques including Southern (DNA) and/or northern (mRNA) hybridization, and amplification techniques such as polymerase chain reaction (PCR), ligase chain reaction (LCR), and the like. Examples of related polynucleotides include human and non-human tank2 genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to TANK2 and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of TANK2.

The invention includes both naturally occurring and non-naturally occurring tankyrase2 polynucleotides and polypeptide products thereof. Naturally occurring tankyrase2 products include distinct polynucleotide and polypeptide tankyrase2 species as they occur in humans. However, the invention includes other human tankyrase2 polynucleotide and polypeptide species defined through the analysis of sequence homology. The invention further comprises corresponding homologs of human TANK2 polypeptides and tank2 polynucleotides that are expressed in cells of other animal species, preferably mammalian homologs, and more preferably primate homologs. Within each tankyrase2 species, the invention further provides splice variants, which are encoded by the same genomic DNA but arise from distinct mRNA transcripts. Non-naturally occurring tankyrase2 products include variants of the naturally occurring tankyrase2 products such as polynucleotide and polypeptide

analogs (i.e., wherein one or more nucleotides or amino acids are added, substituted, or deleted). Non-naturally-occurring TANK2 polypeptide products also include TANK2 products that have been covalently modified, e.g., water-soluble polymer modifications, glycosylation variants, and the like.

The tankyrase2 polypeptides and the nucleic acids that encode the polypeptides provide a basis for diagnostic methods for the precise and accurate detection and/or quantitation of TANK2 expression and medical conditions associated with excessive or insufficient TANK2 activity. Furthermore, the nucleotide sequences disclosed herein may be used in the detection of aberrations, such as mutations and deletions, in the gene encoding TANK2. For example, the nucleotide sequences disclosed herein may be used to identify and isolate a genomic sequence for tank2. PCR primers can be designed from various portions of the introns and exons of a genomic tank2 nucleic acid sequence that will allow detection of aberrations in the genomic sequence.

The invention further provides methods of using TANK2 and genetically engineered host cells that express recombinant TANK2 to evaluate and screen for modulators of the poly(ADP-ribose) polymerase activity of the enzyme. Such screening methods may be used for the identification of allosteric agonists and antagonists of TANK2 activity as well as for the identification of direct (e.g., competitive inhibitors) of such activity. TANK2 protein antagonists and inhibitors, such as anti-TANK2 antibodies and tank2 antisense molecules, will provide the basis for pharmaceutical compositions for the treatment and amelioration of symptoms associated with excessive poly(ADP-ribose) polymerase activity. Agonists of TANK2 will provide the basis of the treatment and amelioration of symptoms associated with insufficient poly(ADP-ribose) polymerase activity.

Tankyrase2 Polynucleotides

The present invention provides, *inter alia*, novel purified and isolated polynucleotides encoding human TANK2 polypeptides. The polynucleotides of the invention include DNA sequences and RNA transcripts, both sense and complementary antisense strands, and splice variants thereof. DNA sequences of the

invention include, without limitation, cDNA and genomic sequences. As used herein, lower case "tank2" refers to a tankyrase2 nucleic acid sequence whereas upper case "TANK2" refers to a tankyrase2 amino acid sequence.

"Nucleic acid" as used herein refers to an oligonucleotide or polynucleotide sequence, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin, which may be double-stranded or single-stranded, whether representing the sense or antisense strand. An exemplary double-stranded polynucleotide according to the invention can have a first strand (i.e., a coding strand) having a sequence encoding a TANK2 polypeptide, along with a second strand (i.e., a "complementary" or "non-coding" strand) having a sequence deducible from the first strand according to the Watson-Crick base-pairing rules for DNA. Double-stranded or "duplex" structures may be DNA:DNA, DNA:RNA, or RNA:RNA nucleic acids. A preferred double-stranded polynucleotide is a cDNA comprising the coding region of a nucleotide sequence defined by SEQ ID NO:132 or SEQ ID NO:134. An exemplary single-stranded polynucleotide according to the invention is a messenger RNA (mRNA) encoding a TANK2 polypeptide. Another exemplary single-stranded polynucleotide is an oligonucleotide probe or primer that hybridizes to the coding or non-coding strand of a polynucleotide selected from among the sequences defined by SEQ ID NO:132, and SEQ ID NO:134. Other alternative nucleic acid structures, e.g., triplex structures, are also contemplated.

Genomic DNA of the invention comprises the protein-coding region for a TANK2 polypeptide and includes allelic variants of the preferred polynucleotides of the invention, such as single nucleotide polymorphisms. Genomic DNA of the invention is distinguishable from genomic DNAs encoding polypeptides other than TANK2 in that it includes the TANK2-coding region found in tank2 cDNA of the invention. Genomic DNA can be transcribed into RNA, and the resulting RNA transcript may undergo one or more splicing events wherein one or more introns (i.e., non-coding regions) of the transcript are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms and therefore be subjected to removal of different non-coding RNA sequences but still encode a TANK2 polypeptide, are referred to in the art as "splice variants," and are embraced by the invention. Splice

variants comprehended by the invention, therefore, are encoded by the same DNA sequences but give rise to different amino acid sequences. Such splice variants can comprise regions in which the reading frame is shifted, wherein a downstream portion of the RNA sequence is translated differently, to yield different amino acid sequences in the resulting polypeptides. Allelic variants are known in the art to be modified forms of the wild-type (predominant) gene sequence. Such modifications result from recombination during chromosomal segregation or exposure to conditions that give rise to genetic mutation. Allelic variants, like wild-type genes, are naturally occurring sequences, as opposed to non-naturally occurring variants, which arise from *in vitro* manipulation.

The invention also comprehends cDNA, which is obtained through reverse transcription of an RNA polynucleotide encoding TANK2 followed by second strand synthesis of a complementary strand to provide a double stranded DNA. For example, the invention provides a cDNA sequence that encodes a polypeptide having an amino acid sequence selected from among the sequences defined by SEQ ID NO:133 and SEQ ID NO:135. In a preferred embodiment, the invention provides polynucleotides comprising the coding region of a nucleotide sequence selected from among the sequences defined by SEQ ID NO:132 and SEQ ID NO:134.

As noted, highly preferred nucleic acid sequences according to the invention are defined by SEQ ID NO:132 or SEQ ID NO:134. However, because the genetic code is redundant or "degenerate" in its information-encoding properties, different nucleotide sequences may encode the same polypeptide sequence. Accordingly, the invention comprises the alternative (degenerate) nucleotide sequences that encode TANK2 polypeptides of the invention and functional equivalents thereof. For example, the invention includes polynucleotides comprising nucleotide sequences that are substantially homologous to the TANK2-encoding regions of the nucleotide sequences set forth in SEQ ID NO:132 or SEQ ID NO:134. More particularly, the invention includes polynucleotides whose corresponding nucleotide sequences have at least 90%, preferably at least 95%, more preferably at least 98%, and still more preferably at least 99% identity with a nucleotide sequence defined in SEQ ID NO:132 or SEQ ID NO:134.

Variant polynucleotides of the invention further include fragments of the tank2 nucleotide sequences defined in SEQ ID NO:132 and SEQ ID NO:134, and homologs thereof. The disclosure of full-length polynucleotides encoding TANK2 polypeptides makes readily available to the person having ordinary skill in the art every possible fragment of the full-length polynucleotides. Preferably, fragment polynucleotides of the invention comprise sequences unique to the TANK2-coding nucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (i.e., specifically) to polynucleotides encoding TANK2 or fragments thereof containing the unique sequence. Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other untranslated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of computer software routinely used in the art, e.g., alignment programs available in public sequence databases.

The invention also provides fragment polynucleotides that are conserved in one or more polynucleotides encoding members of the TANK2 family of polypeptides. Such fragments include sequences characteristic of the family of TANK2 polypeptides, referred to as "signature" sequences. The conserved signature sequences are readily discernable following simple sequence comparison of polynucleotides encoding members of the TANK2 family. Polynucleotide fragments of the invention can be labeled in a manner that permits their detection, including radioactive and non-radioactive labeling.

Hybridization can be defined to include the process of forming partially or completely double-stranded nucleic acid molecules through sequence-specific association of complementary single-stranded nucleic molecules. The invention, therefore, further encompasses the use of nucleic acid species that hybridize to the coding or non-coding strands of a polynucleotide that encodes a TANK2 protein. Preferred hybridizing species hybridize to the coding or non-coding strand of the nucleotide sequence defined by SEQ ID NO:132 or SEQ ID NO:134. Also encompassed are species that would hybridize to a TANK2-encoding polynucleotide

but for the redundancy of the genetic code, i.e., polynucleotides that encode the same amino acid sequence but rely on different codon usage.

Hybridizing species include, for example, nucleic acid hybridization or amplification probes (oligonucleotides) that are capable of detecting nucleotide sequences (e.g., genomic sequences) encoding TANK2 or closely related molecules, such as alleles. The specificity of the probe, i.e., whether it is derived from a highly conserved, conserved, or non-conserved region or domain, and the stringency of the hybridization or amplification conditions (high, intermediate, or low) will determine whether the probe identifies only naturally occurring tank2, or related sequences. Probes for the detection of related nucleotide sequences are selected from conserved or highly conserved regions of tank2 family members and such probes may be used in a pool of degenerate probes. For the detection of identical nucleotide sequences, or where maximum specificity is desired, oligonucleotide probes are selected from the non-conserved nucleotide regions or unique regions of tank2 polynucleotides. As used herein, the term "non-conserved nucleotide region" refers to a nucleotide region that is unique to tank2 disclosed herein and does not occur in related tank2 family members.

Specificity of hybridization is typically characterized in terms of the degree of stringency of the conditions under which the hybridization is performed. The degree of stringency of hybridization conditions can refer to the melting temperature (T_m) of the nucleic acid binding complex [see, e.g., Berger and Kimmel, "Guide to Molecular Cloning Techniques," *Methods in Enzymology*, Vol. 152, Academic Press, San Diego, CA (1987)]. "Maximal stringency" typically occurs at about T_m – 5°C (5°C below the T_m of the probe); "high stringency" at about 5°C to 10°C below T_m; "intermediate stringency" at about 10°C to 20°C below T_m; and "low stringency" at about 20°C to 25°C below T_m.

Alternatively, the stringency of hybridization can refer to the physicochemical conditions employed in the procedure. To illustrate, exemplary moderately stringent hybridization conditions are: hybridization in 3X saline sodium citrate (SSC), 0.1% sarkosyl, and 20 mM sodium phosphate, pH 6.8, at 65°C; and washing in 2X SSC with 0.1% sodium dodecyl sulfate (SDS), at 65°C. Exemplary highly stringent

hybridization conditions are: hybridization in 50% formamide, 5X SSC, at 42°C overnight, and washing in 0.5X SSC and 0.1% SDS, at 50°C. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described Ausubel et al. (Eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons (1994), at pp. 6.0.3-6.4.10. Modifications in hybridization conditions can be determined empirically or calculated precisely based on the length of the oligonucleotide probe and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook et al., (Eds.), *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47-9.51.

The artisan will appreciate that hybridization under more stringent conditions enables the identification of species having a higher degree of homology or sequence identity with the target sequence. By contrast, hybridization under less stringent conditions enables identification of species having a lesser but still significant degree of homology or sequence identity with the target sequence. Therefore, also included within the scope of the present invention are nucleic acid species that are capable of hybridizing to the nucleotide sequence of SEQ ID NO:132 or SEQ ID NO:134 under conditions of intermediate (moderate) to maximal stringency. Preferably, the hybridizing species hybridize to the coding or non-coding strands of a polynucleotide defined by SEQ ID NO:132 or SEQ ID NO:134 under highly stringent conditions.

The polynucleotides of the invention encompass oligonucleotides (i.e., nucleic acid oligomers typically about 10 to 60 nucleotides in length) that hybridize to either the coding or the non-coding strands of a nucleic acid encoding a TANK2 amino acid sequence. In particular, the invention comprises oligonucleotides that hybridize to the coding or non-coding strand of a polynucleotide defined by SEQ ID NO:132 or SEQ ID NO:134. The length of the oligonucleotide is not critical, as long as it is capable of hybridizing to the target nucleic acid molecule. However, longer nucleic acid molecules are more difficult to prepare and require longer hybridization times. Therefore, the oligonucleotide should not be longer than necessary. Accordingly, the oligonucleotide should contain at least 10 nucleotides, preferably at least 15

nucleotides, and more preferably at least 20 nucleotides. Normally, the oligonucleotide will not contain more than 60 nucleotides, preferably not more than 30 nucleotides, and more preferably not more than 25 nucleotides. Such oligonucleotides may be used as described herein as primers for DNA synthesis (e.g., as primers in PCR; "amplimers"), as probes for detecting the presence of target DNA in a sample (e.g., northern or Southern blots and *in situ* hybridization), as therapeutic agents (e.g., in antisense therapy), or for other purposes. Oligonucleotides may be single- or double-stranded, with the double-stranded forms having one or both ends blunt or stepped.

The oligonucleotides may be obtained or derived by known methods from natural sources. Alternatively, the oligonucleotides may be produced synthetically according to methods known in the art. Such methods include, for example, cloning and restriction of appropriate sequences or direct chemical synthesis by any suitable method. Various chemical methods for making oligonucleotides are known in the art, including the phosphotriester method, the phosphodiester method; the diethylphosphoramidite method; the solid support method, and the H-phosphonate method [for reviews, see Caruthers, *Science* 230:281-5 (1985); Caruthers et al., *Methods Enzymol* 211:3-20 (1992)]. Typically, preparation of oligonucleotides is carried out by automated phosphoramidite synthesis on polymer support. Nucleic acid molecules consisting of 100 or more nucleotides may also be produced by such methods.

The tank2 polynucleotides of the invention include variants, which are polynucleotides that encode hAPRP2 or a functional equivalent thereof, and which can include deletions, insertions, or substitutions of nucleotide residues. As used herein a "deletion" is a change in a nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent. As used herein an "insertion" or "addition" is a change in a nucleotide or amino acid sequence that results in the addition of one or more nucleotides or amino acid residues, respectively. As used herein a "substitution" is a change in a nucleotide or amino acid sequence in which one or more nucleotides or amino acids are replaced by different nucleotides or amino acids, respectively.

Polynucleotide variants also included within the scope of the present invention are alleles or alternative naturally occurring forms of tank2. Alleles result from naturally occurring mutations, i.e., deletions, insertions or substitutions, in the genomic nucleotide sequence, which may or may not alter the structure or function or the expressed polypeptides. Each of these types of mutational changes may occur alone, or in combination with the others, one or more times in a given allelic sequence. Single nucleotide polymorphisms (SNPs) may occur, in which a single base mutation may define an altered polypeptide, which in turn may be associated with an overt phenotypic difference. Of course, SNPs may be silent, as they may not change the encoded polypeptide, or any change they do encode may have no effect on phenotype.

The invention further embraces natural homologs of the human tankyrase2 DNA that occur in other animal species, such as other mammal species. Mammalian homologs include, for example, homologs in mouse, rat, guinea pig, and the like, and more preferably homologs in other primate species. Such species homologs, in general, share significant homology at the nucleotide level within the protein-coding regions. Thus, the invention encompasses polynucleotides that share at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% nucleotide identity with the protein-coding region of a polynucleotide encoding a human TANK2 polypeptide, e.g., a polynucleotide defined by SEQ ID NO:132 or SEQ ID NO:134. Percent sequence "homology" with respect to polynucleotides of the invention can be defined as the percentage of nucleotide bases in a candidate sequence that are identical to nucleotides in the TANK2-encoding sequence after aligning the sequences and introducing gaps, if necessary, to achieve maximum percent sequence identity. Computer software is available (from commercial and public domain sources) for calculating percent identity in an automated fashion (e.g., FASTA).

The invention includes polynucleotides that have been engineered to selectively modify the cloning, processing, and/or expression of the TANK2 gene product. Mutations may be introduced using techniques well known in the art, e.g., site-directed mutagenesis to insert new restriction sites, to alter glycosylation patterns,

or to change codon preferences inherent in the use of certain expression systems, while simultaneously maintaining control of the amino acid sequence of the expressed polypeptide product. For example, codons preferred by a particular prokaryotic or eukaryotic host cell can be selected ("codon optimization") to increase the rate of TANK2 expression or to produce recombinant RNA transcripts having desirable properties, such as longer half-lives.

The tank2 polynucleotides can be synthesized, wholly or partly, using chemical methods well known in the art. "Chemically synthesized," as used herein and is understood in the art, refers to purely chemical, as opposed to enzymatic, methods for producing polynucleotides. "Wholly" chemically synthesized DNA sequences are therefore produced entirely by chemical means; "partly" chemically synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means.

DNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences of the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiester linkages within the backbone of the molecule.

The invention also provides TANK2 peptide nucleic acid (PNA) molecules. These TANK2 PNAs are informational molecules that have a neutral "peptide-like" backbone with nucleobases that allow the molecules to hybridize to complementary TANK2-encoding DNA or RNA with higher affinity and specificity than corresponding oligonucleotides (PerSeptive Biosystems).

Polypeptide Expression Systems

Knowledge of TANK2-encoding DNA sequences enables the artisan to modify cells to permit or increase expression of TANK2. Accordingly, host cells are provided, including prokaryotic or eukaryotic cells, either stably or transiently modified by introduction of a polynucleotide of the invention to permit expression of the encoded TANK2 polypeptide. Autonomously replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating TANK2-encoding sequences are also provided.

Expression constructs are also provided comprising TANK2-encoding polynucleotides operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, and operators, and are generally selected based on the expression systems in which the expression construct is to be used. Preferred promoter and enhancer sequences are generally selected for the ability to increase gene expression, while operator sequences are generally selected for the ability to regulate gene expression. Preferred constructs of the invention also include sequences necessary for replication in a host cell. Expression constructs are preferably used for production of an encoded TANK2 polypeptide, but may also be used to amplify the construct itself.

Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA in to a host cell include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include, for example, bacteria, yeast, fungal, plant, insect, invertebrate, amphibian, and mammalian cell systems. Some suitable prokaryotic host cells include, for example, *E. coli* strains SG-936, HB 101, W3110, X1776, X2282, DHI, and MRC1, *Pseudomonas* sp., *Bacillus* sp. such as *B. subtilis*, and *Streptomyces* sp. Suitable eukaryotic host cells include yeasts, such as *Saccharomyces cerevisiae*, *S. pombe*, *Pichia pastoris* and other fungi, insect cells such as sf9 or sf21 cells (*Spodoptera frugiperda*), animal cells such as Chinese hamster ovary (CHO) cells, human cells such as JY, 293, and NIH3T3 cells, and plant cells such as *Arabidopsis thaliana* cells. The tank2 nucleotide sequence, or any portion of it, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6.

The type of host cell, the form of the expressed TANK2 product, the conditions of growth, etc., can be selected by the skilled artisan according to known

criteria. Use of mammalian host cells is expected to provide for such post-translational modifications (e.g., glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of TANK2 polypeptides are embraced. The protein produced by a recombinant cell may be secreted or may be contained intracellularly, depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing tank2 can be designed with signal sequences that direct secretion of TANK2 through a particular prokaryotic or eukaryotic cell membrane.

Expression constructs may include sequences that facilitate, and preferably promote, homologous recombination in a host cell. This can be accomplished by replacing all or part of the naturally occurring tank2 promoter with all or part of a heterologous promoter so that the cells express TANK2 at higher levels. The heterologous promoter should be inserted so that it is operatively linked to TANK2-encoding sequences. See, for example, PCT International Publication Nos. WO 94/12650, WO 92/20808, and WO 91/09955.

Host cells of the invention are useful in methods for large-scale production of TANK2 polypeptide products. For example, host cells of the invention are a valuable source of immunogen for development of antibodies that are immunoreactive with TANK2 polypeptides. As another example, recombinant TANK2 can be produced and isolate from host cells for use in *in vitro* binding assays such as drug screening assays. In such methods, the host cells are grown in a suitable culture medium and the desired polypeptide product is isolated from the cells or from the medium in which the cells are grown.

The polypeptide product can be isolated by purification methods known in the art, such as conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high performance liquid chromatography (HPLC), reverse phase HPLC, and the like.

Still other methods of purification include those in which the desired protein is expressed and purified as a fusion protein in which the TANK2 polypeptide is ligated to a heterologous amino acid sequence. Suitable heterologous sequences can include a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. For example, for screening of peptide libraries for modulators of TANK2 activity, it is possible to express a TANK2 protein fused to a selected heterologous protein selected to be specifically identifiable using a probe antibody. A fusion protein may also be engineered to contain a cleavage site (e.g., a factor XA or enterokinase sensitive sequence) located between the TANK2 sequence and the heterologous protein sequence, to permit the TANK2 protein to be cleaved from the heterologous protein and subsequently purified. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues resulting from the cleavage process.

Exemplary heterologous peptide domains include metal-chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals [Porath, *Protein Expr Purif* 3:263-81 (1992)], and protein A domains that allow purification on immobilized immunoglobulin. Another useful system is the divalent cation-binding domain and antibodies specific thereto used in the peptide extension/immunoaffinity purification system described in US Patents Nos. 4,703,004; 4,782,137; 4,851,431; and 5,011,912. This system is commercially available as the FLAG® system from Immunex Corp. (Seattle WA). Another suitable heterologous fusion partner is glutathione *S*-transferase (GST), which can be affinity purified using immobilized glutathione. Other useful fusion partners include immunoglobulins and fragments thereof, e.g., Fc fragments.

Identification of host cells expressing recombinant TANK2 may be crucial to identifying appropriate expression systems. Accordingly, expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct in operative condition. It is also contemplated that, in addition to the insertion of heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene that encodes carbamyl phosphate synthase, aspartate transcarbamylase, and

dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the TANK2-encoding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the TANK2-encoding sequences in the cells. Detection of expression of the marker gene in response to induction or selection usually indicates expression of TANK2 as well. Alternatively, if the tank2 polynucleotide is inserted within a marker gene sequence, recombinant cells containing tank2 can be identified by the absence of marker gene function.

Host cells that contain the coding sequence for TANK2 and express TANK2 may also be identified by a variety of other procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridization and protein bioassay or immunoassay techniques that include membrane-based, solution-based, or chip-based technologies for the detection and/or quantification of the nucleic acid or protein.

The presence of the tank2 polynucleotide sequence can be detected by DNA-DNA or DNA-RNA hybridization or amplification using fragments of a tank2 polynucleotide, e.g., fragments of the sequences set forth in SEQ ID NO:132 or SEQ ID NO:134, as probes. Nucleic acid amplification based assays involve the use of oligonucleotides based on the tank2 sequence to detect transformants containing tank2 DNA or RNA. Labeled hybridization or PCR probes for detecting tank2 polynucleotide sequences can be made by various methods, including oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide.

In an embodiment of the present invention, TANK2 or a variant thereof and/or a host cell line that expresses the TANK2 or variant thereof may be used to screen for antibodies, peptides, or other molecules, such as organic or inorganic molecules, that act as modulators of a biological or immunological activity of TANK2. For example, anti-TANK2 antibodies capable of neutralizing the polymerase or DNA-binding activity of TANK2 may be used to inhibit TANK2-mediated cell death. Alternatively, screening of peptide libraries or organic libraries made by combinatorial chemistry with recombinantly expressed TANK2 or variants thereof or cell lines expressing TANK2 or variants thereof may be useful for identification of therapeutic molecules

that function by modulating a biological or immunological activity of TANK2. Synthetic compounds, natural products, and other sources of potentially biologically active materials can be screened in a number of ways deemed routine by those of skill in the art. For example, nucleotide sequences encoding the DNA-binding domain of TANK2 may be expressed in a host cell, which can be used for screening of allosteric modulators, either agonists or antagonists, of TANK2 activity. Alternatively, nucleotide sequences encoding the conserved catalytic domain of TANK2 can be expressed in host cells and used to screen for inhibitors of ADP-ribose polymerization.

TANK2 Polypeptides

The invention also provides purified and isolated mammalian TANK2 polypeptides. Exemplary TANK2 polypeptides have amino acid sequences defined in SEQ ID NO:133 or SEQ ID NO:135. TANK2 polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. TANK2 products of the invention may be full-length polypeptides, or variant polypeptide products such as fragments, truncates, deletion mutants, and other variants thereof that retain specific TANK2 biological activity. As used herein, "biologically active" refers to a TANK2 polypeptide having structural, regulatory or biochemical functions of the naturally occurring TANK2 protein. Specifically, a TANK2 protein of the present invention has the ability to bind DNA and to polymerize ADP-ribose subunits in response to DNA damage in a cell.

The protein and fragments of the present invention may be prepared by methods known in the art. Such methods include isolating the protein directly from cells, isolating or synthesizing DNA encoding the protein and using the DNA to produce recombinant protein, and synthesizing the protein chemically from individual amino acids.

The TANK2 polypeptides can be isolated from a biological sample, such as a solubilized cell fraction, by standard methods. Some suitable methods include precipitation and liquid chromatographic protocols such as ion exchange, hydrophobic

interaction, and gel filtration [see, e.g., Deutscher (Ed.), *Methods Enzymol* (*Guide to Protein Chemistry*, Section VII) 182:309 (1990) and Scopes, *Protein Purification*, Springer-Verlag, New York (1987)]. Alternatively, purified material is obtained by separating the protein on preparative SDS-PAGE gels, slicing out the band of interest and electroeluting the protein from the polyacrylamide matrix by methods known in the art. The detergent SDS is removed from the protein by known methods, such as by dialysis or the use of a suitable column, such as the Extracti-Gel® column from Pierce Chemical Co. (Rockford, IL).

The TANK2 polypeptide of the invention may also be chemically synthesized, wholly or partly, by methods known in the art [see, e.g., Stuart and Young, *Solid Phase Peptide Synthesis*, 2d ed., Pierce Chemical Co. (1984)]. For example, peptides can be synthesized by solid phase techniques, cleaved from the resin, and purified by preparative HPLC [see, e.g., Roberge et al., *Science* 269:202-4 (1995)]. Automated synthesis may be accomplished, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Norwalk, CT) in accordance with the instructions provided by the manufacturer. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure).

Recombinant TANK2 protein may be produced in and isolated from a host cell transformed with an expression vector containing a tank2 nucleotide sequence and grown in cell culture. As described herein, the host cells, either prokaryotic or eukaryotic, are either stably or transiently transfected (eukaryotic) or transformed (prokaryotic) with a TANK2-encoding polynucleotide of the invention in manner that permits directed expression of a TANK2 polypeptide. In such methods, the host cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells or from the medium in which the cells are grown. Isolation of the polypeptides can be accomplished by, for example, immunoaffinity purification. The use of transformed host cells is preferred for large-scale production of TANK2 polypeptides.

The invention includes polypeptides comprising amino acid sequences that are substantially homologous to the sequences of TANK2 polypeptides described herein. For example, the invention includes polypeptides whose corresponding amino acid

sequences have at least 90%, preferably at least 95%, more preferably at least 98%, and still more preferably at least 99% identity with the polypeptide sequence defined in SEQ ID NO:133 or SEQ ID NO:135.

Percent sequence "identity" with respect to a preferred polypeptide of the invention can be defined as the percentage of amino acid residues in a candidate sequence that are identical to amino acid residues in the reference TANK2 sequence after aligning the sequences and introducing gaps, if necessary, to achieve maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity.

Percent sequence "homology" with respect to a preferred polypeptide of the invention can be defined as the percentage of amino acid residues in a candidate sequence that are identical to amino acid residues in the reference TANK2 sequence after aligning the sequences and introducing gaps, if necessary, to achieve maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity.

Determinations of whether two amino acid sequences are substantially homologous can also be based on FASTA searches [Pearson et al., *Proc Natl Acad Sci USA* 85:2444-8 (1988)]. Alternatively, percent homology is calculated as the percentage of amino acid residues in the smaller of the two sequences that align with identical amino acid residues in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment [see Dayhoff, in *Atlas of Protein Sequence and Structure*, Vol. 5, National Biochemical Research Foundation, Washington, D.C. (1972), at p. 124].

A polypeptide may be considered homologous to a TANK2 polypeptide of the invention if polynucleotides encoding the two polypeptides hybridize with one another. A higher degree of homology is shown if the hybridization occurs under hybridization conditions of greater stringency. Control of hybridization conditions and the relationships between hybridization conditions and degree of homology are understood by those skilled in the art [see, e.g., Sambrook et al., *supra*]. Thus, a

homologous polypeptide may be a polypeptide that is encoded by a polynucleotide that hybridizes with a polynucleotide encoding a polypeptide of the invention under hybridization conditions having a specified degree of stringency.

It may be desirable that such structurally homologous polypeptides will also exhibit functional homology, insofar as the homologous polypeptide has substantially the same function as the polypeptide of the invention. For example, structurally homologous polypeptides may be considered functionally homologous if they exhibit similar binding of a ligand, or similar immune reactivity, etc.

However, it is known that two polypeptides or two polynucleotides may be considered to be substantially homologous in structure, and yet differ substantially in function. For example, single nucleotide polymorphisms (SNPs) among alleles may be expressed as polypeptides having substantial differences in function along one or more measurable parameters such as antibody- or ligand-binding affinity or enzymatic substrate specificity, and the like. Other structural differences, such as substitutions, deletions, splicing variants, and the like, may affect the function of otherwise structurally identical or homologous polypeptides.

The TANK2 polypeptides of the invention include functional derivatives of a TANK2 polypeptides defined in SEQ ID NO:133 or SEQ ID NO:135. Such functional derivatives include polypeptide products that possesses a structural feature or a biological activity that is substantially similar to a structural feature or a biological activity of the TANK2 protein. Accordingly, functional derivatives include variants, fragments, and chemical derivatives of the parent TANK2 protein.

As used herein "variant" refers to a molecule substantially similar in structure and function to either the entire TANK2 molecule, or to a fragment thereof. A molecule is said to be "substantially similar" to another, if both molecules have substantially similar structures or if both molecules possess a similar biological activity. Thus, provided that two molecules possess a similar activity, they are considered variants, as that term is used herein, even if one of the molecules possesses a structure not found in the other molecule, or if the sequence of amino acid residues is not identical.

Among the variant polypeptides provided under the invention are variants that

comprise one or more changes in the amino acid sequence of the TANK2 polypeptide. Such sequence-based changes include deletions, substitutions, or insertions in the TANK2 sequence, as well as combinations thereof.

Deletion variants of the TANK2 polypeptides are polypeptides in which at least one amino acid residue of the sequence is removed. Deletions can be effected at one or both termini of the protein, or with removal of one or more residues within the TANK2 amino acid sequence. Deletion variants include, for example, all incomplete fragments of the TANK2 polypeptides of the invention. As used herein "fragment" refers to any polypeptide subset of the TANK2 protein.

Fragments of TANK2 that exhibit a biological activity characteristic of TANK2 and that are soluble (i.e., not membrane bound) are desirable. A soluble fragment is preferably generated by deleting any membrane-spanning region(s) of the parent molecule or by deleting or substituting hydrophilic amino acid residues for hydrophobic residues. Identification of such residues is well known in the art.

Substitution variants are provided, including polypeptides in which at least one amino acid residue of a TANK2 polypeptide is replaced by an alternative residue. Any substitution can be made, with conservative substitutions being preferred. Directed amino acid substitutions may be made based on well defined physicochemical parameters of the canonical and other amino acids (e.g., the size, shape, polarity, charge, hydrogen-bonding capacity, solubility, chemical reactivity, hydrophobicity, hydrophilicity, or the amphipathic character of the residues.) as well as their contribution to secondary and tertiary protein structure. Substitution variants can include polypeptides comprising one or more conservative amino acid substitutions, i.e., a substitution of one amino acid by another having similar physicochemical character as desired. To illustrate, the canonical amino acids can be grouped according to the following categories:

Aliphatic Side Chains	Gly, Ala; Val, Leu, Ile
Aromatic Side Chains	Phe, Tyr, Trp
Aliphatic Hydroxyl Side Chains	Ser, Thr
Basic Side Chains	Lys, Arg, His
Acidic Side Chains	Asp, Glu
Amide Side Chains	Asn, Gln
Sulfur-Containing Side Chains	Cys, Met
Secondary Amino Group	Pro

Substitutions are preferably made in accordance with the following Table 1 when it is desired to controllably define the characteristics of the TANK2 molecule.

TABLE 1

Original Residue	Exemplary Conservative Substitutions
Ala	gly; ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	ala; pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; tyr; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

Substantial changes in functional or immunological identity are made by selecting substitutions that are more progressive than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. The substitutions that are in general more progressive are those in which: (a) glycine and/or proline is substituted by another amino acid or is deleted or inserted; (b) a hydrophilic residue is substituted for a hydrophobic residue; (c) a cysteine residue is substituted for (or by) any other residue; (d) a residue having an electropositive side chain is substituted for (or by) a residue having an electronegative charge; or (e) a residue having a bulky side chain is substituted for (or

by) one not having such a side chain. Most preferred are amino acid substitutions that affect the solubility of TANK2. These are most preferably generated by substituting hydrophilic for hydrophobic amino acids.

Substitution variants, however, can include non-canonical or non-naturally occurring amino acid residues substituted for amino acid residues in the principal sequence. Substitution variants include those polypeptides in which amino acid substitutions have been introduced by modification of polynucleotides encoding a TANK2 polypeptide.

Insertion variants are provided, in which at least one amino acid residue is present in addition to a TANK2 amino acid sequence. Insertions may be located at either or both termini of the polypeptide, or may be positioned within the TANK2 amino acid sequence. Insertional variants also include fusion proteins in which the amino or carboxy terminus of the TANK2 polypeptide is fused to another polypeptide. Examples of such fusion proteins include immunogenic polypeptides, proteins with long circulating half-life (e.g., immunoglobulin constant regions), marker proteins (e.g., green fluorescent protein) and proteins or polypeptides that facilitate purification of the desired TANK2 polypeptide (e.g., FLAG® tags or polyhistidine sequences). Another example of a terminal insertion is a fusion of a signal sequence, whether heterologous or homologous to the host cell, to the N-terminus of the molecule to facilitate the secretion of the derivative from recombinant hosts. Intrasequence insertions (i.e., insertions within a TANK2 molecule sequence) may range generally from about 1 to 10 residues, more preferably 1 to 5.

Polypeptide variants of the invention also include mature TANK2 products, i.e., TANK2 products wherein leader or signal sequences are removed, as well as products having additional amino terminal residues. TANK2 products having an additional methionine residue at position -1 (Met^{-1}-TANK2) are contemplated, as are TANK2 products having additional methionine and lysine residues at positions -2 and -1, respectively (Met^{-2}-Lys^{-1}-TANK2). Other such variants are particularly useful for recombinant protein production in bacterial host cells.

The invention also encompasses TANK2 variants having additional amino acid residues resulting from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as a glutathione-S-transferase (GST) fusion product yields the desired polypeptide having an additional glycine residue at position -1 (Gly^{-1}-TANK2) upon cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

The invention further provides TANK2 polypeptide products that are chemical derivatives of a TANK2 polypeptide defined in SEQ ID NO:133 or SEQ ID NO:135. As used herein, the term "chemical derivative" refers to molecules that contain additional chemical moieties that are not normally a part of the naturally occurring molecule. Such moieties may impart desirable properties to the derivative molecule, such as increased solubility, absorption, biological half-life, etc. The moieties may alternatively decrease the toxicity of the derivative molecule, or eliminate or attenuate any undesirable side effect of the derivative molecule. Thus, chemical derivatives of TANK2 polypeptides include polypeptides bearing modifications other than (or in addition to) insertion, deletion or substitution of amino acid residues. Preferably, the modifications are covalent in nature, and include, for example, chemical bonding with polymers, lipids, non-naturally occurring amino acids, and other organic and inorganic moieties. Derivatives of the invention may be prepared to increase circulating half-life of a TANK2 polypeptide, or may be designed to improve targeting capacity for the polypeptide to desired cells, tissues, or organs.

For example, methods are known in the art for modifying a polypeptide to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Particularly preferred are TANK2 products that have been covalently modified with polyethylene glycol (PEG) subunits. Water-soluble polymers may be bonded at specific positions, for example at the amino terminus of the TANK2 products, or randomly attached to one or more side chains of the polypeptide. Additional derivatives include TANK2 species

immobilized on a solid support, pin microparticle, or chromatographic resin, as well as TANK2 species modified to include one or more detectable labels, tags, chelating agents, and the like.

Derivatization with bifunctional agents can be used to cross-link TANK2 to a water-insoluble support matrix. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates and reactive substrates may be employed for protein immobilization [see, e.g., US Patent Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440.]

Expression of TANK2 variants can be expected to have utility in investigating a biological activity characteristic of a wild-type TANK2 polypeptide. TANK2 variants can be designed to retain all biological or immunological properties characteristic for TANK2, or to specifically disable one or more particular biological or immunological properties of TANK2. For example, fragments and truncates may be designed to delete a domain associated with a particular property, or substitutions and deletions may be designed to inactivate a property associated with a particular domain. Forced expression (overexpression) of such variants ("dominant negative" mutants) can be employed to study the function of the protein *in vivo* by observing the phenotype associated with the mutant.

Functional derivatives of TANK2 having up to about 100 residues may be conveniently prepared by *in vitro* synthesis. If desired, such fragments may be modified using methods known in the art by reacting targeted amino acid residues of the purified or crude protein with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues. The resulting covalent derivatives may be used to identify residues important for biological activity.

Functional derivatives of TANK2 having altered amino acid sequences can also be prepared by mutating the DNA encoding TANK2. Any combination of amino acid deletion, insertion, and substitution may be employed to generate the final

construct, provided that the final construct possesses the desired activity. Obviously, the mutations that will be made in the DNA encoding the functional derivative must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure [see EP Patent Publication No. 75,444].

While the site for introducing a variation in the amino acid sequence is predetermined, the mutation per se need not be predetermined. For example, to optimize the performance of a mutation at a given site, random mutagenesis, such as linker scanning mutagenesis, may be conducted at a target codon or target region to create a large number of derivative which could then be expressed and screened for the optimal combination of desired activity. Alternatively, site-directed mutagenesis or other well-known technique may be employed to make mutations at predetermined sites in a DNA known sequence.

The technique of site-directed mutagenesis is well known in the art [see, e.g., Sambrook et al., *supra*, and McPherson (Ed.), *Directed Mutagenesis: A Practical Approach*, IRL Press, Oxford (1991)]. Site-directed mutagenesis allows the production of TANK2 functional derivatives through use of specific oligonucleotide sequences that encode the DNA sequence of the desired mutation. Site-directed mutagenesis methods and materials are commercially available, e.g., the QuikChange™ kit available from Stratagene (La Jolla, CA). One can selectively generate precise amino acid deletions, insertions, or substitutions using this method. Amino acid sequence deletions generally range from about 1 to 30 residues, more preferably 1 to 10 residues, and typically are contiguous. The most preferred deletions are those that are performed to generate catalytic fragments or DNA-binding fragments.

Mutations designed to increase the affinity of TANK2 may be guided by the introduction of the amino acid residues that are present at homologous positions in other poly(ADP-ribose) polymerase proteins. Similarly, such mutant TANK2 molecules may be prepared that lack residues of a functional domain, e.g., the catalytic domain, to create a dominant negative protein.

It is difficult to predict *a priori* the exact effect any particular modification, e.g., substitution, deletion, insertion, etc., will have on the biological activity of TANK2. However, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. For example, a derivative typically is made by linker scanning site-directed mutagenesis of the DNA encoding the native TANK2 molecule. The derivative is then expressed in a recombinant host, and, optionally, purified from the cell culture, for example, by immunoaffinity chromatography. The activity of the cell lysate or the purified derivative is then screened in a suitable screening assay for the desired characteristic. For example, a change in the immunological character of the functional derivative, such as affinity for a given antibody, is measured by a competitive type immunoassay. Changes in other parameters of the expressed product may be measured by the appropriate assay.

Antibodies

The present invention provides antibodies that bind with specificity to a TANK2 polypeptide. An "antibody" as used herein is defined broadly as a protein that characteristically immunoreacts with an epitope (antigenic determinant) that is characteristic of the TANK2 polypeptide. As used herein, an antibody is said to "immunoreact" with an antigen such as a polypeptide if the antibody specifically recognizes and binds an epitope that is characteristic of the antigen by way of one or more variable regions or one or more of the complementarity determining regions (CDRs) of the antibody.

An antibody that is immunoreactive with a given polypeptide may exhibit cross-reactivity to another polypeptide if the two polypeptides each comprise a common structural feature that defines the same characteristic epitope. In the case of related polypeptides, cross-reactivity can correlate to common structural features such as sequence identity, homology, or similarity found among the related polypeptides. Accordingly, families of polypeptides can often be identified by a cross-reactive antibody, i.e., an antibody that immunoreacts with some or all of the members of the polypeptide family sharing the common epitope. Thus, the invention encompasses antibodies that immunoreact with a particular member of the TANK2 family of

polypeptides, e.g., a polypeptide comprising the amino acid sequence defined by SEQ ID NO:133 or SEQ ID NO:135. The invention further encompasses antibodies that immunoreact with some or all members of the TANK2 family of polypeptides. Screening assays to determine the binding specificity of an antibody are well known and routinely practiced in the art [see, e.g., Harlow et al. (Eds.), *Antibodies: A Laboratory Manual*, Ch. 6, Cold Spring Harbor Laboratory, Cold Spring Harbor NY (1988)]. The immunoreactive specificity with which an antibody binds to a given polypeptide antigen is to be distinguished from interactions with other proteins, e.g., *Staphylococcus aureus* protein A or other antibodies in ELISA techniques, that are mediated through parts of the antibody other than the variable regions, in particular the constant regions of the antibody.

Antibodies include, for example, monoclonal antibodies, polyclonal antibodies, single chain antibodies (scFv antibodies), chimeric antibodies, multifunctional/multispecific (e.g., bifunctional or bispecific) antibodies, humanized antibodies, human antibodies, and CDR-grafted antibodies (including moieties that include CDR sequences that specifically immunoreact with a polypeptide of the invention). Antibodies according to the invention also include antibody fragments, so long as they exhibit the desired biological activity. "Antibody fragments" comprise a portion of a full-length antibody, generally the antigen binding or variable region thereof. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and Fv fragments; diabodies; linear antibodies; single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Antibodies of the invention can be produced by any method known in the art. For example, polyclonal antibodies are isolated from mammals that have been immunized against the protein or a functional analog in accordance with methods known in the art. Briefly, polyclonal antibodies may be produced by injecting an immunogenic TANK2 polypeptide (immunogen) into a host mammal (e.g., rabbit, mouse, rat, or goat). Adjuvants may be employed to increase the immune response. Sera from the host mammal are extracted and screened to obtain polyclonal antibodies that are specific for (immunoreact with) the TANK2 polypeptide.

Monoclonal antibodies (also referred to herein as "mAbs") are preferred. As used herein "monoclonal antibody" refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific ("monospecific"), being directed against a single antigenic site. Furthermore, in contrast to conventional (polyclonal) antibody preparations, which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen.

The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. Monoclonal antibodies may be prepared using any suitable technique capable of yielding a continuous cell line producing a homogeneous antibody. Such methods include the immunological method [Köhler and Milstein, *Nature* 256:495-7 (1975); Campbell, "Monoclonal antibody technology, the production and characterization of rodent and human hybridomas" in Burdon et al. (Eds.), *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 13, Elsevier Science Publishers, Amsterdam (1985)] or any similar method. Monoclonal antibodies may also be isolated from phage antibody libraries [Clackson et al., *Nature* 352:624-8 (1991); Marks et al., *J Mol Biol* 222:581-97 (1991)].

To illustrate, to produce monoclonal antibodies a host mammal is immunized by injection of an immunogenic TANK2 polypeptide, and then boosted. Spleens are collected from immunized mammals a few days after the final boost. Cell suspensions from the spleens are fused with a tumor cell line to create immortalized hybrid cell lines or "hybridomas." Individual clones can be isolated by limiting dilution and then tested for the specificity of the antibodies they produce. Selected cells can then be grown, e.g., by the ascites method, to provide a continuous source of the desired homogeneous antibody.

Antibodies can be engineered using genetic techniques to produce chimeric antibodies including protein components from two or more species. For use in *in vivo* applications with a human subject, the antibody can be "humanized," i.e., modified to contain an antigen binding region from one species, e.g., a rodent, with the bulk of the antibody replaced with sequences derived from human immunoglobulin. In one method, the non-human CDRs of one species e.g., a mouse or rabbit, are inserted into a framework sequence of another species, e.g., a human, or into a consensus framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity of the engineered antibody. Methods are also known for inducing expression of engineered antibodies in various cell types, such as mammalian and microbial cell types. Numerous techniques for preparing engineered antibodies are described in the art [e.g., Owens and Young, *J Immunol Meth* 168:149-65 (1994)].

Antibodies further include recombinant polyclonal or monoclonal Fab fragments [e.g., Huse et al., *Science* 246:1275-81 (1989)]. Alternatively, techniques described for the production of single chain antibodies [e.g., US Patent No. 4,946,778] can be adapted to produce TANK2-specific single chain antibodies (e.g., single chain Fv fragments; abbreviated "scFv"). Rapid, large-scale recombinant methods for generating antibodies may be employed, such as phage display or ribosome display methods, optionally followed by affinity maturation [see, e.g., Ouwehand et al., *Vox Sang* 74(Suppl 2):223-32 (1998); Rader et al., *Proc Natl Acad Sci USA* 95:8910-5 (1998); Dall'Acqua et al., *Curr Opin Struct Biol* 8:443-50 (1998)].

Fully human antibodies are especially preferred for therapeutic use in humans, but they are typically difficult to produce. For example, when the immunogen is a human self-antigen, a human will typically not produce any immune response to the antigen. Methods for making fully human antibodies have been developed and are known in the art. Accordingly, fully human antibodies can be produced by using an immunogenic TANK2 polypeptide to immunize an animal (e.g., mouse) that has been transgenically modified to express at least a significant fraction of the human repertoire of immunoglobulin genes [see, e.g., Bruggemann et al., *Immunol Today* 17:391-7 (1996)].

As noted herein, host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with TANK2. To be useful as an immunogen for the preparation of polyclonal or monoclonal antibodies, a TANK2 peptide fragment must contain sufficient amino acid residues to define an immunogenic epitope. If the fragment is too short to be immunogenic *per se*, it may be conjugated to a carrier molecule. Suitable carrier molecules include, for example, keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Conjugation may be carried out by methods known in the art. One such method is to combine a cysteine residue of the fragment with a cysteine residue on the carrier molecule.

Antibodies of the invention are useful for therapeutic methods (by modulating activity of TANK2), diagnostic methods (by detecting TANK2 in a sample), as well as purification of TANK2. The antibodies are particularly useful for detecting and/or quantitating TANK2 expression in cells, tissues, organs, and lysates and extracts thereof, as well as in fluids such as serum, plasma, cerebrospinal fluid, urine, sputum, peritoneal fluid, pleural fluid, or bronchoalveolar lavage fluid. Kits comprising an antibody of the invention for any of the purposes described herein are also contemplated. In general, a kit of the invention also includes a control antigen with which the antibody immunoreacts, and may further include other reagents, containers, and package inserts.

Further, the invention includes neutralizing antibodies, i.e., antibodies that significantly inhibit or impair a biological activity of the proteins or functional analogs of the invention. In particular, neutralizing antibodies inhibit or impair the poly(ADP-ribose) polymerase activity of TANK2. Neutralizing antibodies may be especially desirable for therapeutic and diagnostic applications.

Functional equivalents further include fragments of antibodies that have the same binding characteristics as, or that have binding characteristics comparable to, those of the whole antibody. Such fragments may contain one or both Fab fragments or the $F(ab')_2$ fragment. Preferably, the antibody fragments contain all six complement determining regions ("CDRs") of the whole antibody, although fragments containing fewer than all of such regions, such as three, four, or five CDRs,

may also be functional. Fragments may be prepared by methods described in the art [e.g., Lamoyi et al., *J Immunol Meth* 56:235-43 (1983); Parham, *J Immunol* 131:2895-902 (1983)].

Moreover, specific binding proteins can be developed using isolated or recombinant TANK2 products, TANK2 variants, or cells expressing such products. Binding proteins are useful for purifying TANK2 products and detection or quantification of TANK2 products in fluid and tissue samples using known immunological procedures. Binding proteins are also manifestly useful in modulating (i.e., blocking, inhibiting, or stimulating) biological activities of TANK2 polypeptides, especially those activities involved in signal transduction. Thus, neutralizing antibodies that inhibit the activity of TANK2 polypeptides are provided. Anti-idiotypic antibodies specific for anti-TANK2 antibodies are also contemplated.

Detectable Polynucleotide and Polypeptide Probes

The present invention further provides a method of detecting the presence of a TANK2-encoding polynucleotide or a TANK2 polypeptide in a sample. The method involves use of a labeled probe that recognizes the presence of a defined target in the sample. The probe may be an antibody that recognizes a TANK2 polypeptide, or an oligonucleotide that recognizes a polynucleotide encoding TANK2 polypeptide.

The probes of the invention can be detectably labeled in accordance with methods known in the art. In general, the probe can be modified by attachment of a detectable label (reporter) moiety to the probe, or a detectable probe can be manufactured with a detectable label moiety incorporated therein. The detectable label moiety can be any detectable moiety, many of which are known in the art, including radioactive atoms, electron dense atoms, enzymes, chromogens and colored compounds, fluorogens and fluorescent compounds, members of specific binding pairs, and the like.

Methods for labeling oligonucleotide probes have been described in the art [see, e.g., Leary et al., *Proc Natl Acad Sci USA* 80:4045-49 (1983); Renz and Kurz, *Nucleic Acids Res* 12:3435-44 (1984); Richardson and Gumport, *Nucleic Acids Res* 11:6167-84 (1983); Smith et al., *Nucleic Acids Res* 13:2399-412 (1985); Meinkoth

and Wahl, *Anal Biochem* 138:267-84 (1984); and US Patent Nos. 4,711,955; 4,687,732; 5,241,060; 5,244,787; 5,328,824; 5,580,990; and 5,714,327].

Methods for labeling antibodies have been also been described [see, e.g., Hunter et al., *Nature* 144:495-6 (1962); David et al., *Biochemistry* 13:1014-21 (1974); and US Patent Nos. 3,940,475 and 3,645,090].

The label moiety may be radioactive. Some examples of useful radioactive labels include ^{32}P, ^{125}I, ^{131}I, and ^{3}H. Use of radioactive labels has been described [e.g., UK patent document 2,034,323 and US Patent Nos. 4,358,535 and 4,302,204].

Some examples of non-radioactive labels include enzymes, chromogens, atoms and molecules detectable by electron microscopy, and metal ions detectable by their magnetic properties.

Some useful enzymatic labels include enzymes that cause a detectable change in a substrate. Some useful enzymes (and their substrates) include, for example, horseradish peroxidase (pyrogallol and o-phenylenediamine), beta-galactosidase (fluorescein beta-D-galactopyranoside), and alkaline phosphatase (5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium). The use of enzymatic labels has been described in the art [see, e.g., UK patent document 2,019,404, European patent document EP 63,879, and Rotman, *Proc Natl Acad Sci USA* 47:1981-91 (1961)].

Useful reporter moieties include, for example, fluorescent, phosphorescent, chemiluminescent, and bioluminescent molecules, as well as dyes. Some specific colored or fluorescent compounds useful in the present invention include, for example, fluoresceins, coumarins, rhodamines, Texas red, phycoerythrins, umbelliferones, Luminol®, and the like. Chromogens or fluorogens, i.e., molecules that can be modified (e.g., oxidized) to become colored or fluorescent or to change their color or emission spectra, are also capable of being incorporated into probes to act as reporter moieties under particular conditions.

The label moieties may be conjugated to the probe by methods that are well known in the art. The label moieties may be directly attached through a functional group on the probe. The probe either contains or can be caused to contain such a

functional group. Some examples of suitable functional groups include, for example, amino, carboxyl, sulfhydryl, maleimide, isocyanate, isothiocyanate.

Alternatively, label moieties such as enzymes and chromogens may be conjugated to antibodies or nucleotides by means of coupling agents, such as dialdehydes, carbodiimides, dimaleimides, and the like.

The label moiety may also be conjugated to the probe by means of a ligand attached to the probe by a method described above and a receptor for that ligand attached to the label moiety. Any of the known ligand-receptor binding pair combinations is suitable. Some suitable ligand-receptor pairs include, for example, biotin-avidin or -streptavidin, and antibody-antigen. The biotin-streptavidin combination may be preferred.

Methods of Using Tankyrase2 Polynucleotides and Polypeptides

The scientific value of the information contributed through the disclosures of DNA and amino acid sequences of the present invention is manifest. As one series of examples, knowledge of the sequence of a cDNA for tank2 makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding TANK2 and TANK2 expression control regulatory sequences. DNA/DNA hybridization procedures carried out with DNA sequences of the invention under moderately to highly stringent conditions are also expected to allow the isolation of DNAs encoding allelic variants of TANK2. Similarly, non-human species genes encoding proteins homologous to TANK2 can also be identified by Southern and/or PCR analysis. As an alternative, complementation studies can be useful for identifying other human TANK2 products as well as non-human proteins, and DNAs encoding the proteins, sharing one or more biological properties of TANK2. Oligonucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express TANK2. Polynucleotides of the invention may also be the basis for diagnostic methods useful for identifying a genetic alteration in the tank2 locus that underlies a disease state. For example, the differential expression or activity of TANK2-LONG and TANK2-SHORT may be capable of correlation with particular disease state(s), rendering one

or both forms of TANK2 suitable as diagnostic markers or as therapeutic targets as described herein. Therefore, selective reagents, e.g., oligonucleotides that selectively hybridize to one form of tank2 or antibodies that selectively immunoreact with one form of TANK2, may be especially useful.

Oligonucleotides of the invention, as described herein, may be used in methods to amplify DNA for various purposes. "Amplification" according to the method of the invention refers to any molecular biology technique for detection of trace levels of a specific nucleic acid sequence by exponentially amplifying a template nucleic acid sequence. In particular, suitable amplification techniques include such techniques as the polymerase chain reaction (PCR), the ligase chain reaction (LCR) and variants thereof. PCR is known to be a highly sensitive technique, and is in wide use [see, e.g., Innis et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., San Diego (1990); Dieffenbach and Dveksler, *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Plainview NY (1995); and US Patents Nos. 4,683,195; 4,800,195; and 4,965,188]. The more recently developed LCR technique is known to be highly specific, and is capable of detecting point mutations [see, e.g., Landegren et al., *Science* 241:1077-80 (1988) and Barany et al., *PCR Methods and Applications* 1:5-16 (1991)]. An LCR kit is available from Stratagene. In certain circumstances, it is desirable to couple the PCR and LCR techniques to improve precision of detection. Other amplification techniques may be employed in accordance to the invention.

Oligonucleotide amplification primers are often provided as matched pairs of single-stranded oligonucleotides; one with sense orientation (5' $\rightarrow$ 3') and one with antisense (3' $\leftarrow$ 5') orientation. Such specific primer pairs can be employed under optimized conditions for identification of a specific gene or condition. Alternatively, the same primer pair, nested sets of oligomers, or even a degenerate pool of oligomers, may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Such oligonucleotides can be used in various methods known in the art to extend the specified nucleotide sequences. These methods permit use of a known

sequence to determine unknown adjacent sequence, thereby enabling detection and determination of upstream sequences such as promoters and regulatory elements.

For example, restriction-site polymerase chain reaction is a direct method that uses universal primers to retrieve unknown sequence adjacent to a known locus [see, e.g., Gobinda et al., *PCR Methods Applic* 2:318-22 (1993)]. In this method, genomic DNA is first amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region [Triglia et al., *Nucleic Acids Res* 16:8186 (1988)]. The primers may be designed using Oligo 4.0 (National Biosciences, Inc., Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72°C. This method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intermolecular ligation and used as a PCR template.

Capture PCR is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome (YAC) DNA [Lagerstrom et al., *PCR Methods Applic* 1:111-9 (1991)]. Capture PCR also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before PCR. Walking PCR is a method for targeted gene walking that permits retrieval of unknown sequence [Parker et al., *Nucleic Acids Res* 19:3055-60 (1991)]. The PromoterFinder™ kit (Clontech, Palo Alto, CA) uses PCR, nested primers, and special libraries to "walk in" genomic DNA. This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

Such methods can be used to explore genomic libraries to extend 5' sequence and to obtain endogenous tank2 genomic sequence, including elements such as promoters, introns, operators, enhancers, repressors, and the like. Preferred libraries

for screening for full-length cDNAs are ones that have been size-selected to include larger cDNAs. In addition, randomly primed libraries are preferred in that they will contain more sequences that contain the 5' and upstream regions of genes.

The oligonucleotide probes may also be used for mapping the endogenous genomic sequence. The sequence may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. These include *in situ* hybridization to chromosomal spreads [Venna et al., *Human Chromosomes: A Manual of Basic Technique*, Pergamon Press, New York NY (1988)], flow-sorted chromosomal preparations, or artificial chromosome constructions such as YACs, bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries.

Hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers are invaluable in extending genetic maps. Examples of genetic maps can be found in the art [e.g., Hodgkin et al., *Science* 270:410-4 (1995) and Murray et al., *Science* 265:2049-54 (1994)]. Often the placement of a gene on the chromosome of another mammalian species may reveal associated markers even if the number or arm of a particular human chromosome is not known. Such sequences can be assigned to particular structural features of chromosomes by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. See, e.g., Gatti et al., *Nature* 336:577-80 (1988). The polynucleotides of the invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., between normal, carrier, or affected individuals. Other types of genetic maps can also be developed, e.g., physical maps of the genome based on sequence-tagged sites (STS) [see, e.g., Hudson et al., *Science* 270:1945-54 (1995)].

The DNA sequence information provided by the present invention also makes possible the development, e.g., through homologous recombination or "knock-out" strategies [Capecchi, *Science* 244:1288-92 (1989)], of animals that fail to express

functional TANK2 or that express a variant of TANK2. Such animals are useful as models for studying the *in vivo* activities of TANK2 and modulators thereof.

As described herein, the invention provides antisense nucleic acid sequences that recognize and hybridize to polynucleotides encoding TANK2. Modifications of gene expression can be obtained by designing antisense sequences to the control regions of the tank2 gene, such as the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between -10 and +10 regions of the leader sequence, are preferred. Antisense RNA and DNA molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. The worker of ordinary skill will appreciate that antisense molecules of the invention include those that specifically recognize and hybridize to tank2 DNA (as determined by sequence comparison of tank2 DNA to DNA encoding other known molecules). The antisense molecules of the invention also include those that recognize and hybridize to DNA encoding other members of the TANK2 family of proteins. Antisense polynucleotides that hybridize to multiple DNAs encoding other members of the TANK2 family of proteins are also identifiable through sequence comparison to identify characteristic or signature sequences for the family of TANK2 proteins. Accordingly, such antisense molecules preferably have at least 95%, more preferably at least 98%, and still more preferably at least 99% identity to the target tank2 sequence.

Antisense polynucleotides are particularly relevant to regulating expression of TANK2 by those cells expressing tank2 mRNA. Antisense polynucleotides (preferably 10 to 20 bp oligonucleotides) capable of specifically binding to tank2 expression control sequences or tank2 RNA are introduced into cells, e.g., by a viral vector or a colloidal dispersion system such as a liposome. The antisense oligonucleotide binds to the tank2 target nucleotide sequence in the cell and prevents transcription or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use under the invention. The antisense oligonucleotides may be further

modified by poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' ends [for a recent review of antisense technology, see Delihas et al., *Nat Biotechnol* 15:751-3 (1997)].

The invention further comprises methods to modulate TANK2 expression by means of ribozyme technology [for a review, see Gibson and Shillitoe, *Mol Biotechnol* 7:125-37 (1997)]. Ribozyme technology can be used to inhibit translation of tank2 mRNA in a sequence specific manner through (i) the hybridization of a complementary RNA to a target mRNA and (ii) cleavage of the hybridized mRNA through endonuclease activity inherent to the complementary RNA. Ribozymes can be identified by empirical methods such as using complementary oligonucleotides in ribonuclease protection assays, but more preferably are specifically designed based on scanning the target molecule for accessible ribozyme cleavage sites [Bramlage et al., *Trends Biotechnol* 16:434-8 (1998)]. Delivery of ribozymes to target cells can be accomplished using either exogenous or endogenous delivery techniques well known and practiced in the art. Exogenous can include use of targeting liposomes or direct local injection. Endogenous methods include use of viral vectors and non-viral plasmids.

Ribozymes can specifically modulate expression of TANK2 when designed to be complementary to regions unique to a polynucleotide encoding TANK2. "Specifically modulate," therefore is intended to mean that ribozymes of the invention recognize only a polynucleotide encoding TANK2. Similarly, ribozymes can be designed to modulate expression of all or some of the TANK2 family of proteins. Ribozymes of this type are designed to recognize nucleotide sequences conserved all or some of the polynucleotides encoding the TANK2 family members.

The invention further embraces methods to modulate transcription of tank2 through use of oligonucleotide-directed triple helix formation (also known as Hogeboom base-pairing methodology) [for a review, see Lavrovsky et al., *Biochem Mol Med* 62:11-22 (1997)]. Triple helix formation is accomplished using sequence-specific oligonucleotides that hybridize to double stranded DNA in the major groove as defined in the Watson-Crick model. This triple helix hybridization compromises the ability of the original double helix to open sufficiently for the binding of

polymerases, transcription factors, or regulatory molecules. Preferred target sequences for hybridization include promoter and enhancer regions to permit transcriptional regulation of TANK2 expression. Oligonucleotides that are capable of triple helix formation can alternatively be coupled to DNA damaging agents, which can then be used for site-specific covalent modification of target DNA sequences [see Lavrovsky et al., *supra*].

Both antisense RNA and DNA molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligonucleotides such as solid-phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* or *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

Mutations in a gene that result in loss of normal function of the gene product may underlie TANK2-related disease states. The invention comprehends gene therapy to restore TANK2 activity as indicated in treating those disease states characterized by a deficiency or absence of poly(ADP-ribose) polymerase activity associated with the TANK2 enzyme. Delivery of functional tank2 gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments) [see, e.g., Anderson, *Nature* 392(6679 Suppl):25-30 (1998)]. Alternatively, it is contemplated that in other disease states, preventing the expression or inhibiting the activity of TANK2 will be useful in treating those disease states. Antisense therapy or gene therapy can be applied to negatively regulate the expression of TANK2.

The DNA and amino acid sequence information provided by the present invention also makes possible the systematic analysis of the structure and function of TANK2 proteins. DNA and amino acid sequence information for TANK2 also permits identification of molecules with which a TANK2 polypeptide will interact.

Agents that modulate (i.e., increase, decrease, or block) TANK2 activity may be identified by incubating a putative modulator with TANK2 and determining the effect of the putative modulator on TANK2 activity. The selectivity of a compound that modulates the activity of the TANK2 polypeptide can be evaluated by comparing its activity on the TANK2 to its activity on other proteins.

Numerous methods are amenable to modification by including TANK2 polypeptides or tank2 polynucleotides of the invention, including cell based methods such as dihybrid and trihybrid screens to detect binding partners and split hybrid screens to detect compounds that disrupt complexing of binding partners. Other methods include *in vitro* methods, such as assays in which a TANK2 polypeptide, tank2 polynucleotide, or a binding partner thereof is immobilized, as well as solution assays, are contemplated under the invention. These methods are exemplified by a general approach that includes the steps of contacting a TANK2 polypeptide with a putative binding partner compound, detecting or measuring binding of the TANK2 polypeptide with the compound, and optionally isolating and/or identifying the binding partner compound.

Cell-based assays include methods of screening genomic DNA or cDNA libraries to identify binding partners of TANK2 polypeptides. Exemplary methods include the dihybrid or two-hybrid screen [Fields and Song, *Nature* 340:245-6 (1989); Fields, *Methods: A Companion to Methods in Enzymology* 5:116-24 (1993)] which can be used identify DNAs encoding binding partners. Modifications and variations of the dihybrid assay are described [Colas and Brent, *Trends Biotechnol* 16:355-63 (1998)]. Trihybrid screens can also be employed [Fuller et al., *Biotechniques* 25:85-8, 90-2 (1998)].

Cell-based methods of the invention may be used to identify components in biological pathways that are mediated by TANK2 biological activity. In one aspect, the method is carried out in a host cell containing a soluble TANK2 polypeptide and a soluble form of its binding partner and wherein decreased of increased binding is quantitated through measurement of a binding-dependent phenotypic change in the host cell that is associated with a change in expression of a reporter gene product.

Alternatively, cell-based assays to identify inhibitors of TANK2 polypeptide interaction with a known binding partner may be based on methods such as the split hybrid assay [PCT patent publication WO 98/13502] and variations thereof [PCT patent publication WO 95/20652].

In vitro methods can comprise the steps of (a) contacting an immobilized TANK2 polypeptide with a candidate binding partner compound, and (b) detecting binding of the candidate compound to the TANK2 polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of the TANK2 polypeptide is detected. Immobilization may be accomplished using any of the methods well known in the art, including bonding to a support, beads, or a chromatographic resin, as well as high affinity interactions such as antibody binding or use of an avidin:biotin type system. Detection of binding of the ligands can be accomplished, for example, by (i) using a detectable (e.g., radioactive or fluorescent) label on the ligand that is not immobilized, (ii) using an antibody immunospecific for the non-immobilized ligand, (iii) using a label on the non-immobilized ligand that promotes excitation of a fluorescent support to which the immobilized ligand is bound, as well as other techniques routinely practiced in the art.

In solution assays, methods of the invention comprise the steps of (a) contacting a TANK2 polypeptide with one or more candidate binding partner compounds, and (b) identifying the compounds that bind to the TANK2 polypeptide. Identification of the compounds that bind TANK2 can be achieved by isolating the TANK2:binding partner complex, and separating the TANK2 polypeptide from the binding partner compound. An additional step of characterizing the physical, biological, or biochemical properties of the binding partner compound is also comprehended under the invention. In one approach the TANK2:binding partner complex is isolated using a second binding partner compound (e.g., an antibody or other protein) that interacts with either of the principal ligands in the complex.

Selective modulators may include, for example, antibodies and other proteins or peptides that selectively or specifically bind to a TANK2 polypeptide or a TANK2-encoding polynucleotide, oligonucleotides that selectively or specifically bind to TANK2 polypeptides or TANK2-encoding polynucleotides, and other non-peptide

compounds (e.g., isolated or synthetic organic molecules) that selectively or specifically react with TANK2 polypeptides or TANK2-encoding polynucleotides. Modulators also include compounds as described above but which interact with a specific binding partner of TANK2 polypeptides. Mutant forms of TANK2, such as those that affect the biological activity or cellular location of the wild-type TANK2, are also contemplated under the invention. Presently preferred targets for the development of selective modulators include, for example:

(1) cytoplasmic or transmembrane regions of TANK2 polypeptides that contact other proteins and/or localize TANK2 within a cell, e.g., to telomeres;

(2) extracellular regions of TANK2 polypeptides that bind specific binding partners;

(3) regions of the TANK2 polypeptides that bind substrate, i.e., ADP-ribose;

(4) allosteric regulatory sites of the TANK2 polypeptides;

(5) regions of the TANK2 polypeptides that mediate multimerization;

(6) regions of TANK2 or other proteins (e.g., TRF1 or TRF2) that act as acceptors ADP-ribosylation.

Still other selective modulators include those that recognize particular regulatory or TANK2-encoding nucleotide sequences. Selective and specific modulators of TANK2 activity may be therapeutically useful in treatment of a wide range of diseases and physiological conditions in which aberrant TANK2 activity is involved.

A TANK2-encoding polynucleotide sequence may be used for the diagnosis of diseases resulting from or associated with TANK2 expression or activity. For example, polynucleotide sequences encoding a TANK2 polypeptide (e.g., TANK2-LONG or TANK2-SHORT) may be used in hybridization or PCR assays of biological samples, e.g., samples or extracts of fluids or tissues from biopsies or autopsies, to detect abnormalities in TANK2 expression. Such qualitative or quantitative methods may include Southern or northern analysis, dot blot, or other membrane-based technologies; PCR technologies; dipstick, pin or chip technologies; and ELISA or other multiple-sample format technologies. These types of techniques are well known in the art and have been employed in commercially available diagnostic kits.

Such assays may be tailored to evaluate the efficacy of a particular therapeutic treatment regimen and may be used in animal studies, in clinical trials, or in monitoring the treatment of an individual patient. To provide a basis for the diagnosis of disease, a normal or standard profile for TANK2 expression must be established. This is accomplished by combining a biological sample taken from a normal subject with a tank2 polynucleotide, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained for normal subjects with a dilution series of positive controls run in the same experiment where a known amount of a purified tank2 polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by a disorder or disease related to TANK2 expression. Deviation between standard and subject values establishes the presence of the disease state. If disease is established, an existing therapeutic agent is administered, and treatment profile or values may be generated. The assay may be repeated on a regular basis to evaluate whether the values progress toward or return to the normal or standard pattern. Successive treatment profiles may be used to show the efficacy of treatment over a period of several days or several months.

Anti-TANK2 antibodies are useful for the diagnosis of conditions, disorders, or diseases characterized by or associated with abnormal expression of a TANK2 polypeptide. Diagnostic assays for TANK2 polypeptides include methods that employ a labeled antibody to detect a TANK2 polypeptide in a biological sample such as a body fluid, cells, tissues, sections, or extracts of such materials. The polypeptides and antibodies of the present invention may be used with or without modification. Preferably, the polypeptide or the antibody will be labeled by linking them, either covalently or non-covalently, with a detectable label moiety as described herein.

Antibody-based methods for detecting the presence of TANK2 polypeptides in biological samples are enabled by virtue of the present invention, including assays for differential detection of TANK2-LONG versus TANK2-SHORT. Assays for detecting the presence of proteins with antibodies have been previously described, and follow known formats, such as enzyme-linked immunosorbent assay (ELISA),

radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS) and flow cytometry, western blots, sandwich assays, and the like. These formats are normally based on incubating an antibody with a sample suspected of containing the TANK2 protein and detecting the presence of a complex between the antibody and the protein. The antibody is labeled either before, during, or after the incubation step. The specific concentrations of antibodies, the temperature and time of incubation, as well as other such assay conditions, can be varied, depending upon various factors including the concentration of antigen in the sample, the nature of the sample, etc. Those skilled in the art will be able to determine operative and optimal assay conditions for each determination by employing routine experimentation [see, e.g., Hampton et al., *Serological Methods: A Laboratory Manual*, APS Press, St Paul, MN (1990)].

To provide a basis for the quantitation of TANK2 protein in a sample or for the diagnosis of disease, normal or standard values of TANK2 polypeptide expression must be established. This is accomplished by combining body fluids or cell extracts taken from a normal sample or from normal subjects, either animal or human, with antibody to a TANK2 polypeptide. The amount of standard complex formation may be quantified by comparing it with a dilution series of positive controls where a known amount of antibody is combined with known concentrations of a purified TANK2 polypeptide. Then, standard values obtained from normal samples may be compared with values obtained from samples from test sample, e.g., subjects potentially affected by a disorder or disease related to a TANK2 expression. Deviation between standard and test values establishes the presence of the disease state.

Methods for Identifying Modulators of Tankyrase2 Activity

The TANK2 protein, as well as fragments thereof possessing biological activity can be used for screening putative modulator compounds in any of a variety of drug screening techniques. The term "modulator" as used herein refers to a compound that acts as an agonist or as an antagonist of TANK2 activity. Modulators according to the invention include allosteric modulators of activity as well as inhibitors of activity. An "agonist" of TANK2 is a compound that enhances or

increases the ability of TANK2 to carry out any of its biological functions. An example of such an agonist is an agent that increases the ability of TANK2 to bind to damaged DNA or to polymerize ADP-ribose. An "antagonist" of TANK2 is a compound that diminishes or abolishes the ability of TANK2 to carry out any of its biological functions. An example of such antagonists is an anti-TANK2 antibody.

Accordingly, the invention provides a method for screening a plurality of test compounds for specific binding affinity with a TANK2 polypeptide, comprising providing a plurality of test compounds; combining a TANK2 polypeptide with each of the plurality of test compounds for a time sufficient to allow binding under suitable conditions; and detecting binding of the TANK2 polypeptide to each of the plurality of test compounds, thereby identifying those test compounds that specifically bind the TANK2 polypeptide.

The present invention also provides a method of identifying a modulator of a biological activity of a TANK2 polypeptide, comprising the steps of a) contacting the compound with a TANK2 polypeptide, b) incubating the mixture of step a) with a substrate under conditions suitable for the biological activity, c) measuring the amount of the biological activity; and d) comparing the amount of biological activity of step c) with the amount of biological activity obtained with the TANK2 polypeptide, incubated without the compound, thereby determining whether the compound stimulates or inhibits the biological activity. In one embodiment of the method, the TANK2 polypeptide is a fragment from the non-catalytic region of the TANK2 and provides a method to identify allosteric modulators of TANK2. In another embodiment, the TANK2 polypeptide is a fragment from the catalytic region of TANK2 and provides a method to identify inhibitors of the biological activity. TANK2-LONG and TANK2-SHORT polypeptides or specific fragments thereof may be employed.

Accordingly, the polypeptide employed in such methods may be free in solution, affixed to a solid support, displayed on a cell surface, or located intracellularly. The modulation of activity or the formation of binding complexes between the TANK2 polypeptide and the agent being tested may be measured. TANK2 polypeptides are amenable to biochemical or cell-based high throughput

screening (HTS) assays according to methods known and practiced in the art, including melanophore assay systems to investigate receptor-ligand interactions, yeast-based assay systems, and mammalian cell expression systems [for a review, see Jayawickreme and Kost, *Curr Opin Biotechnol* 8:629-34 (1997)]. Automated and miniaturized HTS assays are also comprehended [e.g., Houston and Banks, *Curr Opin Biotechnol* 8:734-40 (1997)].

Such HTS assays are used to screen libraries of compounds to identify particular compounds that exhibit a desired property. Any library of compounds may be used, including chemical libraries, natural product libraries, combinatorial libraries comprising random or designed oligopeptides, oligonucleotides, or other organic compounds.

Chemical libraries may contain known compounds, proprietary structural analogs of known compounds, or compounds that are identified from natural product screening.

Natural product libraries are collections of materials isolated from naturals sources, typically, microorganisms, animals, plants, or marine organisms. Natural products are isolated from their sources by fermentation of microorganisms followed by isolation and extraction of the fermentation broths or by direct extraction from the microorganisms or tissues (plants or animal) themselves. Natural product libraries include polyketides, non-ribosomal peptides, and variants (including non-naturally occurring variants) thereof [for a review, see Cane et al., *Science* 282:63-8 (1998)].

Combinatorial libraries are composed of large numbers of related compounds, such as peptides, oligonucleotides, or other organic compounds as a mixture. Such compounds are relatively straightforward to design and prepare by traditional automated synthesis protocols, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries.

Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries [for a review of combinatorial chemistry and libraries created thereby, see Myers, *Curr Opin Biotechnol* 8:701-7 (1997)].

Once compounds have been identified that show activity as modulators of TANK2 function, a program of optimization can be undertaken in an effort to improve the potency and or selectivity of the activity. This analysis of structure-activity relationships (SAR) typically involves of iterative series of selective modifications of compound structures and their correlation to biochemical or biological activity. Families of related compounds can be designed that all exhibit the desired activity, with certain members of the family potentially qualifying as therapeutic candidates.

The invention also encompasses the use of competitive drug screening assays in which neutralizing antibodies capable of binding a TANK2 polypeptide specifically compete with a test compound for binding to the TANK2 polypeptide. In this manner, the antibodies can be used to detect the presence of any compound, e.g., another peptide that shares one or more antigenic determinants with the TANK2 polypeptide.

Therapeutic Uses of TANK2-Encoding Polynucleotides and TANK2 Polypeptides

The invention provides a method for inhibiting the expression or activity of TANK2 therapeutically or prophylactically in a human or other animal. The method comprises administering a TANK2 antagonist in an amount effective for inhibiting TANK2 expression or activity. The invention thus provides a method for treating tissue damage resulting from cell damage or death due to necrosis or apoptosis, comprising administering to the animal an effective amount of a compound that inhibits TANK2 activity. This method may be employed in treating animals that are or may be subject to any disorder whose symptoms or pathology is mediated by TANK2 expression or activity. Antagonists having specificity for TANK2-LONG or TANK2-SHORT may have particular utility in diseases whose pathology or symptoms are mediated by a specific form of TANK2.

The method may further involve administering an antagonist of another poly(ADP-ribose) polymerase activity, such as activity associated with the enzymes PARP, tankyrase 1, and the like. Exemplary PARP antagonists suitable for use in this embodiment include, for example, the compounds described by Banasik et al. [*J Biol*

Chem 267:1569-75 (1992)]. Other exemplary compounds include those described in PCT patent publications WO 99/11623 and WO 99/11649. Alternatively, the TANK2 inhibitory method may entail use of a compound that antagonizes both TANK2 and another enzyme having poly(ADP-ribose) polymerase activity.

"Treating" as used herein refers to preventing a disorder from occurring in an animal that may be predisposed to the disorder, but has not yet been diagnosed as having it; inhibiting the disorder, i.e., arresting its development; relieving the disorder, i.e., causing its regression, or ameliorating the disorder, i.e., reducing the severity of symptoms associated with the disorder. "Disorder" is intended to encompass medical disorders, diseases, conditions, syndromes, and the like, without limitation.

The methods of the invention embrace various modes of treating an animal in which TANK2 is expressed, and in which TANK2-mediated disorders may be treated. Animals treatable according to the invention include mammals (including humans) and non-mammalian animals, e.g., birds, fish, reptiles, and amphibians. Among the non-human mammals that may be treated are companion animals (pets) including dogs and cats; farm animals including cattle, horses, sheep, pigs, and goats; laboratory animals including rats, mice, rabbits, guinea pigs, and primates. The method is most preferably employed in the treatment of TANK2-mediated disorders in humans.

In particular, the method of the invention may be employed to treat animals therapeutically or prophylactically who are or may subject to a disorder associated with excessive or undesirable telomerase activity. One aspect of the present invention derives from the ability of TANK2 and its functional derivatives to interact with damaged DNA and to modulate the activity of telomere repeat binding factors (e.g., TRF1 and TRF2).

Excessive telomerase activity in cells has been shown to correlate with induction of apparently unlimited capacity of the cells to replicate. In addition, evidence exists that telomerase activity is higher in tumor tissue than most normal tissues suggesting that increased telomerase activity may be essential for tumor growth. Accordingly, the invention also provides to a method of inhibiting oncogenic transformation or inhibiting neoplastic tissue growth, e.g., cancer, in an animal, comprising administering to the animal an effective amount of a compound that

inhibits TANK2 activity. In this embodiment, the method may further comprise adjuvant administration of a chemotherapeutic or anti-cancer drug and/or radiation therapy.

Tumors or neoplasms include new growths of tissue in which the multiplication of cells is uncontrolled and progressive. Some such growths are benign, but others are termed "malignant," leading to death of the organism. Malignant neoplasms or "cancers" are distinguished from benign growths in that, in addition to exhibiting aggressive cellular proliferation, cancers invade surrounding tissues and metastasize. Moreover, malignant neoplasms are characterized in that they show a greater loss of differentiation (greater "dedifferentiation"), and of their organization relative to one another and their surrounding tissues. This property is also called "anaplasia."

Neoplasms treatable by the present invention include solid tumors, i.e., carcinomas and sarcomas. Carcinomas include those malignant neoplasms derived from epithelial cells which tend to infiltrate (invade) the surrounding tissues and give rise to metastases. Adenocarcinomas are carcinomas derived from glandular tissue or in which the tumor cells form recognizable glandular structures. Another broad category of cancers includes sarcomas, which are tumors whose cells are embedded in a fibrillar or homogeneous substance like embryonic connective tissue. The invention also enables treatment of cancers of the myeloid or lymphoid systems, including leukemias, lymphomas and other cancers that typically do not present as a tumor mass, but are distributed in the vascular or lymphoreticular systems.

The type of cancer or tumor cells amenable to treatment according to the invention include, for example, ACTH-producing tumor, acute lymphocytic leukemia, acute nonlymphocytic leukemia, cancer of the adrenal cortex, bladder cancer, brain cancer, breast cancer, cervical cancer, chronic lymphocytic leukemia, chronic myelocytic leukemia, colorectal cancer, cutaneous T-cell lymphoma, endometrial cancer, esophageal cancer, Ewing's sarcoma, gallbladder cancer, hairy cell leukemia, head and neck cancer, Hodgkin's lymphoma, Kaposi's sarcoma, kidney cancer, liver cancer, lung cancer (small and non-small cell), malignant peritoneal effusion, malignant pleural effusion, melanoma, mesothelioma, multiple myeloma,

neuroblastoma, glioma, non-Hodgkin's lymphoma, osteosarcoma, ovarian cancer, ovarian (germ cell) cancer, pancreatic cancer, penile cancer, prostate cancer, retinoblastoma, skin cancer, soft tissue sarcoma, squamous cell carcinomas, stomach cancer, testicular cancer, thyroid cancer, trophoblastic neoplasms, uterine cancer, vaginal cancer, cancer of the vulva, and Wilm's tumor.

As noted above, regulation of telomere structure appears to be associated with aging. Drugs that modulate the regulation of telomere structure can be expected to have utility in treatment of age-related syndromes or in cases of genetically determined premature aging and premature senility syndromes e.g., progeria (Hutchinson-Gilford progeria syndrome), Werner's syndrome, and other such disorders. Accordingly, the invention provides a method of enhancing the activity of TANK2 in animals suffering from such syndromes. The method may be expected to decrease TRF binding to the telomeres, which in turn promotes increased telomerase activity.

Shortening of telomeres beyond a critical length results in the induction of senescence in many cell types. As telomerase activity is frequently required for maintenance of telomere length, and since TANK2 inhibition may diminish telomerase function, the invention provides for treatment of non-neoplastic proliferative disorders in which TANK2 antagonists may be useful to induce shortened telomeres and cellular senescence. Proliferative disorders include, but are not limited to, andrestenosis, diabetic retinopathy, mesangial proliferative disorder, proliferative glomerulonephritis, polycythemia, myelofibrosis, post-transplantation lymphoproliferative disorder, endometriosis, craniosynostosis, immunoproliferative small intestinal disease, thymic lymphoproliferative disease, myelodysplastic disorders, myeloproliferative disorders, von Willebrand's disease, and proliferative nephritis.

In addition, TANK2 inhibitors may be useful in any inflammatory disorder, including autoimmune disorders, in which proliferation of lymphocytes plays a role. "Inflammatory disorder" as used herein can refer to any disease, disorder, or syndrome in which an excessive or unregulated inflammatory response leads to excessive inflammatory symptoms, host tissue damage, or loss of tissue function.

"Inflammatory disorders" can also refer to pathological states mediated by influx of leukocytes and or neutrophil chemotaxis.

"Inflammation" as used herein refers to a localized, protective response elicited by injury or destruction of tissues, which serves to destroy, dilute or wall off (sequester) both the injurious agent and the injured tissue. Inflammation is notably associated with influx of leukocytes and or neutrophil chemotaxis. Inflammation may result from infection with pathogenic organisms and viruses and from noninfectious means such as trauma or reperfusion following myocardial infarction or stroke, immune response to foreign antigen, and autoimmune responses. Inflammatory disorders amenable to the invention encompass disorders associated with reactions of the specific defense system as well as with reactions of the non-specific defense system.

Accordingly, the present invention enables methods of treating such inflammatory disorders as arthritic diseases, such as rheumatoid arthritis, osteoarthritis, gouty arthritis, spondylitis; Behcet disease; sepsis, septic shock, endotoxic shock, gram negative sepsis, gram positive sepsis, and toxic shock syndrome; multiple organ injury syndrome secondary to septicemia, trauma, or hemorrhage; ophthalmic disorders such as allergic conjunctivitis, vernal conjunctivitis, uveitis, and thyroid-associated ophthalmopathy; eosinophilic granuloma; pulmonary or respiratory disorders such as asthma, chronic bronchitis, allergic rhinitis, ARDS, chronic pulmonary inflammatory disease (e.g., chronic obstructive pulmonary disease), silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis, pneumonia, bronchiectasis, and pulmonary oxygen toxicity; reperfusion injury of the myocardium, brain, or extremities; fibrosis such as cystic fibrosis; keloid formation or scar tissue formation; atherosclerosis; autoimmune diseases such as systemic lupus erythematosus (SLE), autoimmune thyroiditis, multiple sclerosis, some forms of diabetes, and Reynaud's syndrome; and transplant rejection disorders such as GVHD and allograft rejection; chronic glomerulonephritis; inflammatory bowel diseases such as Crohn's disease, ulcerative colitis and necrotizing enterocolitis; inflammatory dermatoses such as contact dermatitis, atopic dermatitis, psoriasis, or urticaria; fever and myalgias due to infection; central or peripheral nervous system

inflammatory disorders such as meningitis, encephalitis, and brain or spinal cord injury due to minor trauma; Sjögren's syndrome; diseases involving leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia; antigen-antibody complex mediated diseases; hypovolemic shock; Type I diabetes mellitus; acute and delayed hypersensitivity; disease states due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte transfusion associated syndromes; and cytokine-induced toxicity.

The tank2 polynucleotides provided by the invention also enable therapeutic applications of these polynucleotides in treating the diseases and disorders described herein whose etiology involves TANK2 expression or activity. For example, a tank2 antisense molecule may provide the basis for treatment of various abnormal conditions related to excessive or undesirable levels of poly(ADP-ribose) polymerase activity. Alternatively, polynucleotide sequences encoding TANK2 may provide the basis for the treatment of various abnormal conditions related to deficiency of poly(ADP-ribose) polymerase activity. Polynucleotides having specificity for one or both of tank2-long and tank2-short may have particular utility in certain diseases.

Expression vectors derived from retroviruses, adenovirus, herpes, or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of recombinant tank2 sense or antisense molecules to the targeted cell population. Methods that are well known to those skilled in the art can be used to construct recombinant vectors containing tank2. See, for example, the techniques described in Sambrook et al., *supra*, and Ausubel et al., *supra*. Alternatively, recombinant tank2 can be delivered to target cells in liposomes.

The cDNA sequence, and/or its regulatory elements, enables researchers to use a tank2 polynucleotide as a tool in sense [Youssoufian and Lodish, *Mol Cell Biol* 13:98-104 (1993)] or antisense [Eguchi et al., *Annu Rev Biochem* 60:631-52 (1991)] investigations of gene function. Oligonucleotides, designed from the cDNA or control sequences obtained from the genomic DNA, can be used *in vitro* or *in vivo* to inhibit expression. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions. Again, tank2-long- or tank2-short-specific sequences may have distinct utilities depending on which form of tank2 is of interest.

Additionally, TANK2 expression can be modulated by transfecting a cell or tissue with expression vectors that express high levels of a tank2 polynucleotide fragment in conditions where it would be preferable to block a biological activity of TANK2. Such constructs can flood cells with untranslatable sense or antisense sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until all copies of the vector are disabled by endogenous nucleases. Such transient expression may be accomplished using a non-replicating vector or a vector incorporating appropriate replication elements.

Methods for introducing vectors into cells or tissue include those methods discussed herein. In addition, several of these transformation or transfection methods are equally suitable for *ex vivo* therapy. Furthermore, the tank2 polynucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including but not limited to such properties as the triplet genetic code and specific base pair interactions.

Pharmaceutical Compositions

The present invention further relates to pharmaceutical compositions that comprise a chemical or biological compound ("agent") that is active as a modulator of TANK2 expression or activity and a biocompatible pharmaceutical carrier, adjuvant, or vehicle. The active agent in the pharmaceutical compositions may be selected from among all or portions of tank2 polynucleotide sequences, tank2 antisense molecules, TANK2 polypeptides, protein, peptide, or organic modulators of TANK2 bioactivity, such as inhibitors, antagonists (including antibodies) or agonists. Preferably, the agent is active in treating a medical condition that is mediated by or characterized by TANK2 expression or activity. The composition can include the agent as the only active moiety or in combination with other nucleotide sequences, polypeptides, drugs, or hormones mixed with excipient(s) or other pharmaceutically acceptable carriers.

Techniques for formulation and administration of pharmaceutical compositions may be found in *Remington's Pharmaceutical Sciences*, 18th Ed., Mack Publishing Co., Easton, PA (1990). The pharmaceutical compositions of the present

invention may be manufactured using any conventional method, e.g., mixing, dissolving, granulating, dragée-making, levigating, emulsifying, encapsulating, entrapping, melt-spinning, spray-drying, or lyophilizing processes. However, the optimal pharmaceutical formulation will be determined by one of skill in the art depending on the route of administration and the desired dosage. Such formulations may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the administered agent. Depending on the condition being treated, these pharmaceutical compositions may be formulated and administered systemically or locally.

The pharmaceutical compositions may be administered to the subject by any conventional method, including parenteral and enteral techniques. Parenteral administration modalities include those in which the composition is administered by a route other than through the gastrointestinal tract, for example, intravenous, intraarterial, intraperitoneal, intramedullary, intramuscular, intraarticular, intrathecal, and intraventricular injections. Enteral administration modalities include, for example, oral (including buccal and sublingual) and rectal administration. Transepithelial administration modalities include, for example, transmucosal administration and transdermal administration. Transmucosal administration includes, for example, enteral administration as well as nasal, inhalation, and deep lung administration; vaginal administration; and rectal administration. Transdermal administration includes passive or active transdermal or transcutaneous modalities, including, for example, patches and iontophoresis devices, as well as topical application of pastes, salves, or ointments. Surgical techniques include implantation of depot (reservoir) compositions, osmotic pumps, and the like. A preferred route of administration for treatment of inflammation would be local or topical delivery for localized inflammation such as arthritis, and intravenous delivery for reperfusion injury or for systemic conditions such as septicemia.

The pharmaceutical compositions are formulated to contain suitable pharmaceutically acceptable carriers, and may optionally comprise excipients and auxiliaries that facilitate processing of the active compounds into preparations that can be used pharmaceutically. The administration modality will generally determine the

nature of the carrier. For example, formulations for parenteral administration may comprise aqueous solutions of the active compounds in water-soluble form. Carriers suitable for parenteral administration can be selected from among saline, buffered saline, dextrose, water, and other physiologically compatible solutions. Preferred carriers for parenteral administration are physiologically compatible buffers such as Hank's solution, Ringer's solutions, or physiologically buffered saline. For tissue or cellular administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art. For preparations comprising proteins, the formulation may include stabilizing materials, such as polyols (e.g., sucrose) and/or surfactants (e.g., nonionic surfactants), and the like.

Alternatively, formulations for parenteral use may comprise suspensions of the active compounds prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, and synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Emulsions, e.g., oil-in-water and water-in-oil dispersions, can also be used, optionally stabilized by an emulsifying agent or dispersant (surface-active materials; surfactants). Liposomes containing the active agent may also be employed for parenteral administration. Aqueous polymers that provide pH-sensitive solubilization and/or sustained release of the active agent may also be used as coatings or matrix structures, e.g., methacrylic polymers such as the Eudragit® series available from Röhm America Inc. (Piscataway, NJ).

Alternatively, the pharmaceutical compositions comprising the agent in dosages suitable for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art. The preparations formulated for oral administration may be in the form of tablets, pills, capsules, cachets, dragées, lozenges, liquids, gels, syrups, slurries, suspensions, or powders. To illustrate,

pharmaceutical preparations for oral use can be obtained by combining the active compounds with a solid excipient, optionally grinding the resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries if desired, to obtain tablets or dragée cores. Note that oral formulations may employ liquid carriers similar in type to those described for parenteral use, e.g., buffered aqueous solutions, suspensions, and the like.

Preferred oral formulations include tablets, dragées, and gelatin capsules. These preparations may contain one or excipients, which include, without limitation:

a) diluents such as sugars, including lactose, dextrose, sucrose, mannitol, or sorbitol;

b) binders such as magnesium aluminum silicate, starch from corn, wheat, rice, potato, etc.;

c) cellulose materials such as methyl cellulose, hydroxypropylmethyl cellulose, and sodium carboxymethyl cellulose, polyvinyl pyrrolidone, gums such as gum arabic and gum tragacanth, and proteins such as gelatin and collagen;

d) disintegrating or solubilizing agents such as cross-linked polyvinyl pyrrolidone, starches, agar, alginic acid or a salt thereof such as sodium alginate, or effervescent compositions;

e) lubricants such as silica, talc, stearic acid or its magnesium or calcium salt, and polyethylene glycol;

f) flavorants, and sweeteners;

g) colorants or pigments, e.g., to identify the product or to characterize the quantity (dosage) of active compound; and

h) other ingredients such as preservatives, stabilizers, swelling agents, emulsifying agents, solution promoters, salts for regulating osmotic pressure, and buffers.

Gelatin capsules include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating such as glycerol or sorbitol. Push-fit capsules can contain the active ingredient(s) mixed with fillers, binders, lubricants, and/or stabilizers, etc. In soft capsules, the active compounds may be dissolved or

suspended in suitable fluids, such as fatty oils, liquid paraffin, or liquid polyethylene glycol with or without stabilizers.

Dragée cores can be provided with suitable coatings such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

The pharmaceutical composition may be provided as a salt of the active agent, which can be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms.

To be effective therapeutically in modulating central nervous system targets, the agents used in the methods of the invention should readily penetrate the blood brain barrier when peripherally administered. Compounds that cannot penetrate the blood brain barrier, however, can still be effectively administered by an intravenous route.

As noted above, the characteristics of the agent itself and the formulation of the agent can influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the administered agent. Such pharmacokinetic and pharmacodynamic information can be collected through pre-clinical *in vitro* and *in vivo* studies, later confirmed in humans during the course of clinical trials. Thus, for any compound used in the method of the invention, a therapeutically effective dose can be estimated initially from biochemical and/or cell-based assays. Then, dosage can be formulated in animal models to achieve a desirable circulating concentration range that modulates TANK2 expression or activity. As human studies are conducted, further information will emerge regarding the appropriate dosage levels and duration of treatment for various diseases and conditions.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the "therapeutic index," which is typically expressed as the ratio

LD_{50}/ED_{50}. Compounds that exhibit large therapeutic indices are preferred. The data obtained from such cell culture assays and additional animal studies can be used in formulating a range of dosage for human use. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity.

For the method of the invention, any effective administration regimen regulating the timing and sequence of doses may be used. Doses of the agent preferably include pharmaceutical dosage units comprising an effective amount of the agent. As used herein, "effective amount" refers to an amount sufficient to modulate TANK2 expression or activity and/or derive a measurable change in a physiological parameter of the subject through administration of one or more of the pharmaceutical dosage units.

Exemplary dosage levels for a human subject are of the order of from about 0.001 milligram of active agent per kilogram body weight (mg/kg) to about 100 mg/kg. Typically, dosage units of the active agent comprise from about 0.01 mg to about 10,000 mg, preferably from about 0.1 mg to about 1,000 mg, depending upon the indication, route of administration, etc. Depending on the route of administration, a suitable dose may be calculated according to body weight, body surface area, or organ size. The final dosage regimen will be determined by the attending physician in view of good medical practice, considering various factors that modify the action of drugs, e.g., the agent's specific activity, the severity of the disease state, the responsiveness of the patient, the age, condition, body weight, sex, and diet of the patient, the severity of any infection, etc. Additional factors that may be taken into account include time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Further refinement of the dosage appropriate for treatment involving any of the formulations mentioned herein is done routinely by the skilled practitioner without undue experimentation, especially in light of the dosage information and assays disclosed, as well as the pharmacokinetic data observed in human clinical trials. Appropriate dosages may be ascertained through use of established assays for determining concentration of the agent in a body fluid or other sample together with dose response data.

The frequency of dosing will depend on the pharmacokinetic parameters of the agent and the route of administration. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Accordingly, the pharmaceutical compositions can be administered in a single dose, multiple discrete doses, continuous infusion, sustained release depots, or combinations thereof, as required to maintain desired minimum level of the agent. Short-acting pharmaceutical compositions (i.e., short half-life) can be administered once a day or more than once a day (e.g., two, three, or four times a day). Long acting pharmaceutical compositions might be administered every 3 to 4 days, every week, or once every two weeks. Pumps, such as subcutaneous, intraperitoneal, or subdural pumps, may be preferred for continuous infusion.

Compositions comprising a compound of the invention formulated in a pharmaceutical acceptable carrier may be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition. Conditions indicated on the label may include treatment of inflammatory disorders, cancer, nervous tissue injury, etc. Kits are also contemplated, wherein the kit comprises a dosage form of a pharmaceutical composition and a package insert containing instructions for use of the composition in treatment of a medical condition.

The following Examples are provided to further aid in understanding the invention. The particular materials and conditions employed are intended to exemplify particular aspects of the invention and should not be construed to limit the reasonable scope thereof.

The Examples presuppose an understanding of conventional methods well-known to those persons having ordinary skill in the art to which the examples pertain, e.g., the construction of vectors and plasmids, the insertion of genes encoding polypeptides into such vectors and plasmids, or the introduction of vectors and plasmids into host cells. Such methods are described in detail in numerous publications including, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989), Ausubel et al. (Eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994); and

Ausubel et al. (Eds.), *Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc. (1999).

EXAMPLE 1

Identification of an EST Related to Human Tankyrase1 and Isolation of a Tankyrase2 Polynucleotide

Using the nucleotide sequence of human tankyrase1 (SEQ ID NO:3) [Smith et al. (1998), *supra*], a search of the National Center for Biotechnology Information (NCBI) Expressed Sequence Tags (EST) database was performed to identify novel genes that are homologous to tankyrase1. The EST database provides 5' and/or 3' nucleotide sequences for cDNA clones from a variety of tissue sources. The NCBI BLASTn program [Altschul et al., *Nucleic Acids Res* 25:3389-402 (1997)] was used to compare the nucleotide query sequence of human tankyrase1 against a nucleotide sequence database and to identify DNA sequences in the EST sequence database that have significant homology to human tankyrase1. This BLASTn search identified two EST sequences of interest: AA307492 (SEQ ID NO:5) cloned from a human colon carcinoma cell line designated HCC, and H17748 (SEQ ID NO:7), cloned from human brain.

A comparison of the AA307492 and tankyrase1 polynucleotides revealed that a region consisting of nucleotides 307 to 432 (nt 307-432) of AA307492 (SEQ ID NO:5) shared significant homology with a region consisting of nt 3313-3438 of tankyrase1 (SEQ ID NO:3); 105 of 126 nucleotides were the same; 83% identity). Nucleotides 307-432 of AA307492 were translated and the predicted protein (SEQ ID NO:6) was compared with tankyrase1 protein (amino acids 1105 to 1146 of SEQ ID NO:4). The proteins were found to be the same at 36 of 42 amino acid positions (86% identity). A comparison of the H17748 and tankyrase1 polynucleotides revealed that nt 3-356 of H17748 (SEQ ID NO:7) shared significant homology with nt 3544-3897 of tankyrase1 (SEQ ID NO:3; 280 of 354 nucleotides were identical; 79% identity). When nt 3-356 of H17748 was translated and the predicted protein (SEQ ID NO:8) was compared with the corresponding region of tankyrase1 (aa 1182-1299 of SEQ ID NO:4), the proteins were found to be the same at 111 of 118 amino acid positions

(94% identity). The putative amino acid sequences of AA307492 and H17748 are homologous to, but distinct from, tankyrase1 protein, indicating that they represented protein products translated from a novel tankyrase gene or genes.

AA307492 and H17748 were used in a search of the GenBank® database using the NCBI UniGene® program in order to identify other EST sequences originating from the same gene(s). The UniGene® program assembles GenBank sequences into a non-redundant set of gene-oriented clusters, with each cluster containing a group of sequences from the same gene. The UniGene® search of the human GenBank® database with AA307492 did not identify any other human EST sequences clustering in the same gene region as AA307492. By contrast, the UniGene® search of the human GenBank database with H17748 identified sixteen human EST sequences belonging in the same gene cluster as H17748, as follows: AA305587 (SEQ ID NO:9), AA371079 (SEQ ID NO:10), AA970617 (SEQ ID NO:11), AI247608 (SEQ ID NO:12), H11505 (SEQ ID NO:13), H11865 (SEQ ID NO:14), H17635 (SEQ ID NO:15), N29528 (SEQ ID NO:16), N57467 (SEQ ID NO:17), R06902 (SEQ ID NO:18), R06946 (SEQ ID NO:19), R14158 (SEQ ID NO:20), R33944 (SEQ ID NO:21), R63031 (SEQ ID NO:22), R63337 (SEQ ID NO:23), and T17118 (SEQ ID NO:24). EST H17748 and EST H17635 contained sequence from opposite ends of the same clone, designated 50806. EST H11505 and EST H11865 contained sequence from opposite ends of the same clone, designated 47912. EST R06902 and EST R06946 contained sequence from opposite ends of the same clone, designated 126654. *E. coli* strains harboring cDNA clones 50806, 47912, and 126654 were purchased from the American Type Culture Collection (ATCC, Rockville, MD), which maintains and makes publicly available deposits of ESTs identified and sequenced by I.M.A.G.E. (Lawrence Livermore National Laboratory, Livermore, CA). The three clones were sequenced as follows:

Clone 50806 was sequenced in its entirety on both strands using primers that hybridized to the vector DNA (SEQ ID NOs:25-26), and primers designed to hybridize to the human cDNA (SEQ ID NOs:27-34).

M13 Forward	TGTAAAACGACGGCCAGT	(SEQ ID NO:25)
M13 Reverse	GGAAACAGCTATGACCATG	(SEQ ID NO:26)
NT-7	TTTGCCGGGTAACCTTGG	(SEQ ID NO:27)
NT-8	CCAAGGTTACCCGGCAAA	(SEQ ID NO:28)
NT-9	GTAGGCCCAGTGTAAATG	(SEQ ID NO:29)
NT-10	CATTTACACTGGGCCTAC	(SEQ ID NO:30)
NT-11	GAGTAAGTTGCAGGGCATGT	(SEQ ID NO:31)
NT-12	ACATGCCCTGCAACTTACTC	(SEQ ID NO:32)
NT-13	GAATCACCGCAGTTACTAAA	(SEQ ID NO:33)
NT-14	TTTAGTAACTGCGGTGATTC	(SEQ ID NO:34)

Clone 47912 was sequenced in its entirety on both strands using primers that hybridized to the vector DNA (SEQ ID NOs:25-26, *supra*), and primers designed to hybridize to the human cDNA (SEQ ID NOs:27-34, *supra*, and SEQ ID NOs:35-37).

NT-15	GGCCTGAAGGTATGGTCGAT	(SEQ ID NO:35)
NT-16	ATCGACCATACCTTCAGGCC	(SEQ ID NO:36)
NT-18	TGAGGGCATTACAGTTTGTT	(SEQ ID NO:37)

Clone 126654 was sequenced in its entirety on both strands using primers that hybridized to the vector DNA: M13 Forward (SEQ ID NO:25, *supra*) and T7 Promoter (SEQ ID NO:38), and primers designed to hybridize to the human cDNA (SEQ ID NOs:27-30, *supra*, and SEQ ID NOs:39-40).

T7 Promoter	TAATACGAACTCACTATAGGG	(SEQ ID NO:38)
NT-5	ATACACTCACCGGAGAAA	(SEQ ID NO:39)
NT-6	TTTCTCCGGTGAGTGTAT	(SEQ ID NO:40)

Upon sequencing, 50806, 47912, and 126654 were found to be consistent with the sequences reported in the EST database. The polynucleotide sequences for 50806, 47912, and 126654 are set out in SEQ ID NOs:41, 43, and 45, respectively. The deduced amino acid sequences for 50806, 47912, and 126654 are set out in SEQ ID NOs:42, 44, and 46, respectively. The sequences of 50806 and 47912 indicated that

the clones were identical, and only 50806 was considered further. 50806 and 126654 contain overlapping nucleotide sequence, but 126654 was 63 base pairs longer at the 5' end, while 50806 was approximately 400 base pairs longer at the 3' end.

50806 was determined to have an open reading (ORF) beginning at nucleotide position 1, a potential intron sequence at nt 358-1138, a stop codon beginning at nt 1999, and a potential poly A tail 474 base pairs 3' to the stop codon. When nt 1-357 of 50806 were compared with nt 3538-3897 of tankyrase1, 283 of 357 nucleotides were the same (79% identical). When 50806 was translated from nt 1-357 and the resultant protein was compared with tankyrase1 (aa 1181-1299), the proteins were the same at 116 of 120 amino acid positions (97% identity).

A putative intron was identified in 50806, consisting of nt 358-1138, which may have been an artifact of cDNA cloning. DNA sequences preceding the putative intron (AG) and at the 3' end of the putative intron (CAG) showed high resemblance to the consensus sequence for exon/intron/exon junctions [Lewin, *GENES IV*, Oxford University Press: New York (1997), at p. 88]. The most common sequence at the 3' end of an exon is AG, and at the 3' end of an intron is CAG. To determine if an intron is included in the 50806 sequence, PCR analysis of genomic DNA is used to verify this prediction.

A comparison of 50806 with tankyrase1 showed that a small region consisting of nt 1139-1198 of 50806 was significantly homologous with nt 3896-3957 of tankyrase1 (40 of 60 nucleotides were the same; 67% identity). When 50806 was translated from nt 1139-1198 and the resultant protein was compared with tankyrase1 (aa 1300 to 1319), the proteins were the same at 14 of 20 amino acid positions (70% identity).

126654 was determined to have an ORF beginning at nucleotide position 1, a stop codon beginning at position 481, and a potential poly A tail 81 base pairs 3' of the stop codon. Comparison of 126654 with tankyrase1 showed that a region consisting of nt 1-480 of 126654 shared significant homology with nt 3478-3957 of tankyrase1 (367 of 481 nucleotides identical; 76% identity). When this region of 126654 was translated and the resultant protein compared with the corresponding region of the tankyrase1 protein (i.e., aa 1160-1319), the proteins were the same at

149 of 160 amino acid positions (97% identity). It is possible that either of the putative poly A tails of 50806 and 126654 were artifacts of cDNA cloning or that 50806 and 126654 represented a population of mRNA that use different polyadenylation sites. 50806 had a stretch of 8 A residues 81 base pairs 3' to the stop codon, indicating that the putative poly A tail of 126654 was most likely a cloning artifact.

Alignment of AA307492 and 126654 with human tankyrase1 using the Sequencher™ program (Gene Codes Corporation, Ann Arbor, MI) suggested that AA307492 was upstream of 126654, and that 11 nucleotides separated AA307492 and 126654. To confirm that AA307492 and 126654 represented polynucleotide sequence from the same gene, a primer (SEQ ID NO:47) corresponding to the sense strand of AA307492 and a primer (SEQ ID NO:48) corresponding to the antisense strand of 126654 were synthesized for use in a polymerase chain reaction (PCR) with human Marathon®-Ready spleen and testis cDNA (Clontech) as the template.

AA307492 sense	CTCCGGACAACAAGGTCTTAACC	(SEQ ID NO:47)
126654 antisense	CCACCTATGTACGCATGCC	(SEQ ID NO:48)

The PCR reaction contained 2.5 μL human spleen Marathon®-Ready cDNA, 2.5 μL human testis Marathon-Ready cDNA, 250 nM each primer, 0.25 mM dNTPs, 1X PCR buffer, 1.8 mM $MgCl_2$, and 5 Units of Taq polymerase (Perkin Elmer). The reaction was performed in a GeneAmp® PCR System 9700 machine (hereinafter "GeneAmp® PCR System 9700"; PE Applied Biosystems, Norwalk CT) and first heated at 94°C for 2 min, followed by 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec, and ended with 7 min at 72°C. The PCR fragment was isolated using gel electrophoresis and a QIAquick® Gel Extraction Kit (hereinafter "QIAquick® kit"; Qiagen, Valencia, CA), according to the manufacturer's instructions. The PCR fragment was directly cloned into pCR®2.1-TOPO® vector (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. The PCR fragment was sequenced with primers that hybridized to the vector DNA (SEQ ID NOs:25 and 26, *supra*), and the sequence of the AA307492/126654 PCR fragment is set out in SEQ ID NO:49. The sequence confirmed that AA307492 was upstream of 126654 and that these two

ESTs were separated by 11 nucleotides, and that AA307492 and 126654 were sequences from a novel gene, designated tankyrase2.

To identify the full-length tankyrase2 gene, a probe was generated from 126654 and used to screen a cDNA library using procedures routinely practiced in the art. 126654 was digested with *Xho*I and *Bgl*II, and an approximately 260 nucleotide fragment designated NT-5' was isolated using gel electrophoresis and the QIAquick® kit. NT-5' was labeled with ^{32}P with a Random Primed DNA Labeling Kit (Boehringer Mannheim/Roche Molecular Biochemicals, Indianapolis, IN) according to the manufacturer's instructions and used to screen 10^6 cDNAs from a human fetal brain library (Stratagene). Hybridization with labeled probe was performed overnight at 65°C in buffer containing: 3X SSC, 0.1% sarkosyl, 20 mM sodium phosphate, pH 6.8, 10X Denhardt's solution, and 50 μg/mL salmon sperm DNA. The filters were washed at 65°C in buffer containing 2X SSC and 0.1% SDS prior to autoradiography. Forty-six positives were obtained with the NT-5' probe, of which fifteen were first characterized with respect to strength of hybridization with NT-5'. Restriction digest mapping and partial sequencing led to the selection of two clones, designated FB2B.1 and FB2D.1, for further characterization.

FB2B.1 was sequenced in its entirety on both strands with primers that hybridized to the vector DNA, including T7 promoter (SEQ ID NO:38, *supra*) and T3 promoter (SEQ ID NO:50), and primers designed to anneal to the cDNA sequence (SEQ ID NOs:51-69).

T3 promoter	ATTTAACCCTCACTAAAGGG	(SEQ ID NO:50)
2B.1 F1	AAAGGCTCCCATCGGCAAAT	(SEQ ID NO:51)
2B.1 F2	GTTGAGGGCATTACAGTTTG	(SEQ ID NO:52)
2B.1 F3	AAAACGTAGAGGCCACTGCT	(SEQ ID NO:53)
2B.1 F4	TGGTGTAGACTGACGCCCTT	(SEQ ID NO:54)
2B.1 F5	TCCGGTGAGTGTATCTTTCC	(SEQ ID NO:55)
2B.1 F6	CTCCTTTGTCTTGGGCATTC	(SEQ ID NO:56)
2B.1 F9	ATCTGCTCTGCCCTCTTGTT	(SEQ ID NO:57)
2B.1 F10	GGGTATCGCGGCAATTTACA	(SEQ ID NO:58)
2B.1 F11	AACAAGAGGGCAGAGCAGAT	(SEQ ID NO:59)

2B.1 F12	TGCCCCATCTCAACTAATAC	(SEQ ID NO:60)
2B.1 R2	GTAATGCCCTCAACAGAACT	(SEQ ID NO:61)
2B.1 R3	GGCGTCAGTCTACACCACTT	(SEQ ID NO:62)
2B.1 R4	TAAATTGCCCGCGATACCCA	(SEQ ID NO:63)
2B.1 R5	CACTCAGTCACTGGTAGGCC	(SEQ ID NO:64)
2B.1 R6	ATCTGCTCTGCCCTCTTGTT	(SEQ ID NO:65)
2B.1 R7	TAGTTGAGATGGGGCACAAG	(SEQ ID NO:66)
2B.1 R8	AAACGTAGAGGCCACTGCTG	(SEQ ID NO:67)
2B.1 R9	CGGGTAACCTTGGGAAAGTC	(SEQ ID NO:68)
2B.1&2D.1	GGGCTTTACTGCTTTACAGA	(SEQ ID NO:69)

FB2D.1 was sequenced in its entirety on both strands with primers that hybridized to the vector DNA (SEQ ID NOs:38 and 50, *supra*) and primers designed to anneal to the cDNA sequence, including 2B.1&2D.1 (SEQ ID NO:69) and SEQ ID NOs:70-87.

2D.1 F1	GTAAGGGCTGCTGACAGTGA	(SEQ ID NO:70)
2D.1 F2	TTACTCCAGCAGAGGGCACT	(SEQ ID NO:71)
2D.1 F3	CTGACGCCCTTCAATGTCTC	(SEQ ID NO:72)
2D.1 F4	GGTACTAAGGCCACAATTCA	(SEQ ID NO:73)
2D.1 F5	GGGTATCGCGGCAATTTACA	(SEQ ID NO:74)
2D.1 F6	GTTGAGGGCATTACAGTTTG	(SEQ ID NO:75)
2D.1 F7	TAACAAGAGGGCAGAGCAGA	(SEQ ID NO:76)
2D.1 F8	AGTTCTGTTGAGGGCATTAC	(SEQ ID NO:77)
2D.1 F9	GGCCTACCAGTGACTGAGTG	(SEQ ID NO:78)
2D.1 F10	GGGCTAGAGGACCTGAAGAG	(SEQ ID NO:79)
2D.1 R2	AGTGCCCTCTGCTGGAGTAA	(SEQ ID NO:80)
2D.1 R3	GGCGTCAGTCTACACCACTT	(SEQ ID NO:81)
2D.1 R4	TGAATTGTGGCCTTAGTACC	(SEQ ID NO:82)
2D.1 R5	ATGCCCAAGACAAAGGAGGA	(SEQ ID NO:83)
2D.1 R6	GTAATGCCCTCAACAGAACT	(SEQ ID NO:84)
2D.1 R7	ATCTGCTCTGCCCTCTTGTT	(SEQ ID NO:85)
2D.1 R8	CGGGTAACCTTGGGAAAGTC	(SEQ ID NO:86)

2D.1 R9 CCGGACAACAAGGTCTTAAC (SEQ ID NO:87).

The polynucleotide sequences for FB2B.1and FB2D.1 are set out in SEQ ID NOs:88 and 90, respectively, and the deduced amino acid sequences of FB2B.1and FB2D.1 are set out in SEQ ID NOs:89 and 91, respectively.

The nucleotide and amino acid sequences of FB2B.1 and tankyrase1 were compared to determine the degree of relatedness between the sequences. A region consisting of nt 4-279 of FB2B.1 (SEQ ID NO:88) was found to have significant identity with nt 1624-1899 of tankyrase1 (SEQ ID NO:3), wherein 203 of 276 nucleotides were identical (73% identity). Nucleotides 402-1254 of FB2B.1 showed significant identity with nt 2022-2874 of tankyrase1, wherein 630 of 853 nucleotides were identical (73% identity). Furthermore, nt 1507-2338 of FB2B.1 showed homology to nt 3112-3943 of tankyrase1, wherein 634 of 832 nucleotides were identical (76% identity). FB2B.1 was determined to have an ORF beginning at nucleotide position 1, a stop codon beginning at position 2353, approximately 1 kb of 3' untranslated sequence, but no apparent poly A tail. A translation of nt 1-2352 of FB2B.1 showed that a region consisting of the predicted amino acid sequence (SEQ ID NO:89) was homologous to a corresponding region of tankyrase1 (aa 540-1327 of SEQ ID NO:4). In this region, the proteins were identical at 623 of 777 amino acid positions (80% identity).

A similar comparison of FB2D.1 was made with tankyrase1. In this case, a region consisting of nt 6-197 of FB2D.1 (SEQ ID NO:90) was significantly related to nt 1708-1899 of tankyrase1, wherein 137 of 192 nucleotides were identical (71% identity). Nucleotides 320-1172 of FB2D.1 were found to share significant homology with corresponding nt 2022-2874 of tankyrase1, wherein 630 of 853 nucleotides were identical (73% identity). Nucleotides 1425-2256 of FB2D.1 showed significant homology with nt 3112-3943 of tankyrase1, wherein 634 of 832 nucleotides were identical (76% identity). FB2D.1 was determined to have an ORF beginning at nucleotide position 3, a stop codon beginning at position 2271, approximately 1.5 kb of 3' untranslated sequence, but no apparent poly A tail. When FB2D.1 was translated (SEQ ID NO:91), a domain predicted by the nt 3-2270 showed homology to

aa 569-1327 of tankyrase1 (SEQ ID NO:4). Here, the proteins were the same at 602 of 749 amino acid positions (80% identity).

FB2B.1 and FB2D.1 were aligned using Sequencher™. FB2B.1 and FB2D.1 contained overlapping polynucleotide sequence, but FB2B.1 was longer at the 5' end by 82 base pairs, and FB2D.1 was longer at the 3' end by approximately 0.5 kb. The nucleotide sequences of FB2B.1 and FB2D.1 were identical in the regions nt 83-2971 of FB2B.1 and nt 1-2889 of FB2D.1. However, the remaining 382 nucleotides of FB2B.1 and 910 nucleotides of FB2D.1 did not align. It is possible that FB2B.1 and FB2D.1 were random primed from different positions in the 3' untranslated region and/or that this misalignment was the result of the presence of a cloning artifact in one or both of the clones. Since FB2B.1 and FB2D.1 did not appear to have poly A tails, the poly A tails of ESTs 50806 and 126654 were most likely cloning artifacts, and the real poly A tail of tankyrase2 was most likely greater than 0.5 kb from the stop codon. A consensus polynucleotide sequence, designated 2B.1/2D.1, was developed from the alignment of FB2B.1 and FB2D.1, and is set out in SEQ ID NO:92. 2B.1/2D.1 contained nt 1-2971 of FB2B.1 and nt 1-2889 of FB2D.1.

Alignment of FB2B.1and FB2D.1 with tankyrase1 using Sequencher™ suggested that neither FB2B.1 nor FB2D.1 represented a full-length gene, and that nucleotide sequence was missing from the 5' end of tankyrase2. Thus, FB2B.1 was digested with *Eco*RI and *Sph*I, and an approximately 466 bp nucleotide fragment located at the immediate 5' end of FB2B.1 (nt 49-515 of SEQ ID NO:88) was isolated using gel electrophoresis and the QIAquick® kit. This fragment was labeled with ^{32}P with a Random Primed DNA Labeling Kit and used as a probe (designated NT-37/38) to screen 10^6 cDNA clones of the fetal brain library (Stratagene) using the conditions and procedures used in the first screening. Fourteen positives were obtained with the NT-37/38 probe, one of which (designated 30B.2A) also hybridized with the NT-5' probe, but which had not been chosen for further characterization at that time. Restriction mapping and partial sequencing led to the selection of 30B.2A for further characterization.

The region of 30B.2A upstream of clone FB2B.1 was sequenced with primers that hybridized to the vector DNA (SEQ ID NOs:38 and 50, *supra*) and primers

designed to anneal to the cDNA sequence, including 2B.1 F4 (SEQ ID NO:54, *supra*) and SEQ ID NOs:93-97).

30B.2A #1	GGGCGGAAAGACGTAGTTGA	(SEQ ID NO:93)
30B.2A #2	GCGGCTGTTCACCTTCTCAG	(SEQ ID NO:94)
30B.2A #5	ACGCAAGTGATGGCAGAAAG	(SEQ ID NO:95)
30B.2A #6	TCACTTGCGTGGCAGTTGAC	(SEQ ID NO:96)
30B.2A #7	GCGGCAGGTTTGTAGATGAC	(SEQ ID NO:97)

The partial polynucleotide sequence of 30B.2A is set out in SEQ ID NO:98, and the partial deduced amino acid sequence is set out in SEQ ID NO:99. Comparison of 30B.2A with the nucleotide sequence of tankyrase1 indicated that significant homology occurred in the region consisting of nt 167-1435 of 30B.2A which corresponded with nt 631-1899 of tankyrase1. In this region, 953 of the 1269 nucleotides were the same (75% identity). 30B.2A was determined to have an ORF beginning at nucleotide position 2. Significant amino acid sequence identity was observed between a 385 amino acid sequence predicted for 30B.2A (based on nt 2-1156) and the corresponding region of tankyrase1 (aa 160-539). In this region, the protein sequences were the same at 319 of 385 amino acid positions (83% identity).

2B.1/2D.1 and 30B.2A were aligned using Sequencher™. 30B.2A 2A contained 1.157 kb of novel sequence before it began overlapping with the 5' end of 2B.1/2D.1, and began overlapping with 2B.1/2D.1 at position 1158. A consensus polynucleotide sequence, designated 2B.1/2D.1/30B.2A, was developed from the alignment of 2B.1/2D.1 and 30B.2A, and is set out in SEQ ID NO:100. 2B.1/2D.1/30B.2A contained nt 1-1157 of 30B.2 and nt 1-2971 of 2B.1/2D.1. The predicted amino acid sequence encoded by nt 2-3508 of SEQ ID NO:100 is set forth as SEQ ID NO:101. The nucleotide sequence of the TANK2-encoding region is set forth as SEQ ID NO:1, and the corresponding TANK2 polypeptide sequence is set forth as SEQ ID NO:2.

EXAMPLE 2

Cloning of 5' End of Tankyrase2

Alignment of 30B.2A with tankyrase1 using the Sequencher™ program suggested that 5' sequence was still lacking from the tankyrase2 gene. To clone the 5' end of human tankyrase2, 5' RACE analysis was performed using a Marathon®-Ready human spleen cDNA library (Clontech) as the template. A primer (NT-Marathon; SEQ ID NO:102) corresponding to the antisense strand of 2B.1/2D.1/30B.2A polynucleotide sequence (nt 337-367 of SEQ ID NO:100) was synthesized for use in a polymerase chain reaction (PCR) with the AP1 primer (Clontech; SEQ ID NO:103) that was designed to anneal to the Marathon® cDNA Adapters ligated to the ends of the cDNAs in the library.

NT-Marathon GAGCATTGGGGTCTGCACCATGTCGCAAAAGG (SEQ ID NO:102)

AP1 CCATCCTAATACGACTCACTATAGGGC (SEQ ID NO:103)

The PCR reaction contained 5 µL human spleen Marathon®-Ready cDNA, 0.20 µM each primer, 0.20 mM dNTPs, 1X Clontech GC 2 PCR buffer, Clontech GC-Melt buffer (0, 0.5, 1.0, or 1.5 M), and 1 µL of Clontech Advantage®-GC 2 polymerase mix. The reactions were performed in a GeneAmp® PCR System 9700 with the following four steps: 1) 1 cycle at 94°C for 1 min; 2) 5 cycles of 94°C for 30 sec and 72°C for 30 sec; 3) 5 cycles of 94°C for 30 sec and 70°C for 30 sec; and 4) 25 cycles of 94°C for 30 sec and 60°C for 30 sec. The reactions were then continued in the GeneAmp® PCR System 9700 under the following conditions: 1) 1 cycle at 94°C for 1 min; 2) 5 cycles of 94°C for 30 sec, and 72°C for 3 min; 3) 5 cycles of 94°C for 30 sec and 70°C for 3 min; and 4) 25 cycles of 94°C for 30 sec and 60°C for 3 min. The PCR fragments were isolated using gel electrophoresis and a QIAquick® kit as directed. The PCR fragments were directly cloned into the pCR®2.1-TOPO® vector, as directed. Because *Taq* polymerase has an error rate of 8.0 x 10^{-6} mutation/ base pair (Cline et al., *Nucleic Acids Res* 24:3546-51), four clones isolated from four separate PCR reactions were sequenced and compared to eliminate the possibility of *Taq* polymerase-induced errors in the 5' RACE sequences. The four 5' RACE clones

were sequenced with the M13 forward and M13 reverse primers (SEQ ID NOs:25 and 26) that hybridize to the vector DNA. The four individual nucleotide sequences were compiled into a consensus nucleotide sequence designated 5'-RACE tank2 that is set out in SEQ ID NO:104, and the deduced amino acid sequence is set out in SEQ ID NO:105. In the consensus nucleotide sequence of 5'-RACE tank2, every base pair was present at the corresponding position in at least three of the four unique clones used to compile the consensus sequence. 5'-RACE tank2 and tankyrase were aligned using the Sequencher™ program. When nt 1-279 of 5'-RACE tank2 (SEQ ID NO:104) were compared with tankyrase no significant similarity was found. 5'-RACE tank2 was determined to have an ORF beginning at nucleotide position 2. When nt 2-277 of 5'-RACE tank2 was translated and the resultant protein was compared with tankyrase, no significant similarity was found.

5'-RACE tank2 and 2B.1/2D.1/30B.2A were aligned using the Sequencher™ program. 5'-RACE tank2 contained 279 bp of novel sequence before it began overlapping with the 5' end of FB2B.1/2D.1/30B.2A, and began overlapping with 2B.1/2D.1/30B.2A at position 280. A consensus polynucleotide sequence designated 2B.1/2D.1/30B.2A/5'-RACE, was developed from the alignment of 5'-RACE tank2 and 2B.1/2D.1/30B.2A and is set out in SEQ ID NO:106. 2B.1/2D.1/30B.2A/5'-RACE contained nt 1-279 of 5'-RACE tank2 and nt 1-4140 of 2B.1/2D.1/30B.2A. The deduced putative amino acid sequence of 2B.1/2D.1/30B.2A/5'-RACE is set out in SEQ ID NO:107.

The presence of a continuous ORF in the 5'-RACE tank sequence suggested that 5' sequence was still lacking from the tankyrase2 gene. Further attempts to obtain additional 5' sequence of tankyrase2 using 5' RACE analysis were unsuccessful. The NCBI BLASTn program was used to compare the nucleotide query sequence of FB2B.1/2D.1/30B.2A against a nucleotide sequence tag database (a non-redundant database of GenBank®+EMBL+DDBJ STS Divisions). This BLASTn search identified a STS tag sequence designated stWI-16054 (GenBank® Accession No. G24639; SEQ ID NO:108). When nt 3608-3985 of 2B.1/2D.1/30B.2A was compared with the antisense complement nt 8-397 of stWI-16054, 361 of 378 nucleotides were the same (96% identical). The Sanger Centre (Cambridge, UK)

Human Genome Clone Search program (**http://www.sanger.ac.uk/cgi-bin/humace/searcher.cgi**) was used to identify BAC clones containing stWI-16054. BAC clone bA329B8 was identified as containing the STS tag stWI-16054. BAC clone bA329B8 originates from the genomic RPCI-11.2 male white blood cell library (Pieter deJong, Roswell Park Cancer Institute, Buffalo, NY) and was purchased from Research Genetics, Inc. (Huntsville, AL). A Large Construct Kit (Qiagen) was used to isolate bA329B8 DNA, which was used as a template in inverse PCR amplification reactions [Ochman et al., "Amplification of Flanking Sequences by Inverse PCR," pp. 219-27 in *PCR Protocols: A Guide to Methods and Applications* (Innis et al., eds.), Academic Press, San Diego, CA (1990)]. The inverse PCR technique allows for the amplification of unknown DNA sequence flanking a region of known sequence. Briefly, template DNA is digested with a restriction enzyme (preferably, one that recognizes a four or five base pair consensus site), followed by circularization of the restriction fragments. Circularized fragments are used as a template in a PCR reaction with two primers designed to anneal to the known flanking sequence but pointed in opposite directions. One microgram (1 μg) of bA329B8 was digested in a 20 μL reaction containing 1X appropriate reaction buffer and 10 units of one of the following restriction enzymes: *Rsa*I (Promega, Madison, WI), *Bfa*I (New England Biolabs, Beverly, MA), or *Tru*9I (Promega). The restriction digests were incubated for one hour at 37°C (*Rsa*I and *Bfa*I) or 65°C (*Tru*9I). The *Rsa*I and *Bfa*I digests were heated at 68°C for 20 minutes to inactivate the restriction enzymes. A QIAquick® kit was used to inactivate the restriction enzyme in the *Tru*9I digest. Ligation reactions contained the following: 20 μL of the *Tru*9I, *Rsa*I, or *Bfa*I reactions, 448 μL distilled water, 50 μL 10X reaction buffer, and 2 μL T4 DNA ligase (5U/μL; Boehringer Mannheim, Indianapolis, IN). Ligations were incubated overnight at 15°C. The DNAs in the ligation reactions were then precipitated by adding 129.26 μL 7 M ammonium acetate and 2.3 mL 95% ethanol. The DNAs were pelleted, washed with 75% ethanol, resuspended in 15 μL distilled water, and used as templates in PCR amplification reactions. A primer (5-Inv-1; SEQ ID NO:109) corresponding to the sense strand of 5'-RACE tank2 (nt 423-443 of SEQ ID NO:104) and a primer (3-Inv-

1; SEQ ID NO:110) corresponding to the antisense strand of 5'-RACE tank2 (nt 364-383 of SEQ ID NO:104) were synthesized for use in PCR amplification reactions.

5-Inv-1	CGCCTGAGAAGGTGAACAGCC	(SEQ ID NO:109)
3-Inv-1	ACGCCTCGAACAGCTCTCGG	(SEQ ID NO:110)

The PCR reactions (final reaction volume of 20 μL) contained 5 μL of the *Tru*9I, *Rsa*I, or *Bfa*I DNA template, 0.20 μM each primer, 0.20 mM dNTPs, 1X Clontech GC 2 PCR buffer, 1.0 M Clontech GC-Melt buffer, and 0.4 μL of Clontech Advantage®-GC 2 polymerase. The reactions were performed in a GeneAmp® PCR System 9700 with the following four steps: 1) 1 cycle at 94°C for 1 minute; 2) 5 cycles of 94°C for 30 seconds and 65°C for 3 minutes and 30 seconds; 3) 5 cycles of 94°C for 30 seconds and 60°C for 3 minutes and 30 seconds; and 4) 25 cycles of 94°C for 30 seconds and 58°C for 3 minutes and 30 seconds. The PCR fragments were isolated using gel electrophoresis and a QIAquick® kit as directed. The PCR fragments were directly cloned into the pCR®2.1-TOPO® vector, as directed. The *Tru*9I, *Rsa*I, and *Bfa*I clones were sequenced with the M13 primers that hybridize to the vector DNA (SEQ ID NOs:25 and 26) and primers designed to anneal to the cDNA sequence (SEQ ID NOs:109-112).

5-Inv-2	GCGTGGGCGCGGCCATGGGACTG	(SEQ ID NO:111)
3-Inv-2	CAGCGCGAATCCGCCGTCCG	(SEQ ID NO:112)

The *Tru*9I, *Rsa*I, and *Bfa*I polynucleotide sequences are set out in SEQ ID NOs:113, 115, and 117, respectively. The deduced amino acid sequences of *Tru*9I, *Rsa*I, and *Bfa*I are set out in SEQ ID NOs:114, 116, and 118, respectively.

Clones *Tru*9I and 5'-RACE tank2 were aligned using the Sequencher™ program. Clone *Tru*9I (SEQ ID NO:113) contained 235 bp of novel sequence before it began overlapping with the 5' end of 5'-RACE tank2 (SEQ ID NO:104), and began overlapping with 5'-RACE tank2 at position 236. When nt 1-235 of clone *Tru*9I were compared with tankyrase no significant similarity was found. Clone *Tru*9I was determined to have an ORF beginning at nucleotide position 3. When clone *Tru*9I was translated from nt 3-236 and the resultant protein was compared with tankyrase no significant similarity was found.

Clone *Rsa*I and 5'-RACE tank2 were aligned using the Sequencher™ program. Clone *Rsa*I (SEQ ID NO:115) contained 654 bp of novel sequence before it began overlapping with the 5' end of 5'-RACE tank2 (SEQ ID NO:104), and began overlapping with 5'-RACE tank2 at position 655. When nt 1-654 of clone *Rsa*I were compared with tankyrase no significant similarity was found. Clone *Rsa*I was determined to have an ORF beginning at nucleotide position 160, with a putative ATG start codon beginning at nucleotide 287. When clone *Rsa*I was translated from nt 287-655 and the resultant protein was compared with tankyrase no significant similarity was found.

Clone *Bfa*I (SEQ ID NO:117) and 5'-RACE tank2 were aligned using the Sequencher™ program. Clone *Bfa*I contained 88 bp of novel sequence before it began overlapping with the 5' end of 5'-RACE tank2 (SEQ ID NO:104), and began overlapping with 5'-RACE tank2 at position 89. When nt 1-88 of clone *Bfa*I were compared with tankyrase no significant similarity was found. Clone *Bfa*I was determined to have an ORF beginning at nucleotide position 3. When clone *Bfa*I was translated from nt 3-89 and the resultant protein compared with tankyrase no significant similarity was found.

To confirm the new polynucleotide sequence obtained from the *Tru*9I, *Rsa*I, and *Bfa*I clones and to determine if introns are present in the new sequence, PCR amplification of cDNA was performed. A primer (5-RSA-1; SEQ ID NO:119) corresponding to the sense strand of clone *Rsa*I (nt 59-84 of SEQ ID NO:115) and a primer (3-Inv-1; SEQ ID NO:110) corresponding to the antisense strand of clone *Rsa*I (nt 708-727 of SEQ ID NO:115) were synthesized for use in PCR amplification reactions.

5-RSA-1 GTTCCTCTAATCAATCCTGAGC (SEQ ID NO:119)

Six separate PCR reactions were performed (designated 18, 19, 20, 24, 25, and 26) to aid in the identification of *Taq* polymerase-induced errors as described above. Each 20 μL reaction contained 5 μL of human spleen, placenta, or testis Clontech Marathon®-Ready cDNA DNA template, 0.20 μM each primer, 0.20 mM dNTPs, 1X Clontech GC 2 PCR buffer, 1.0 M Clontech GC-Melt buffer, and 0.4 μL of Clontech Advantage®-GC 2 polymerase. The reactions were performed in a GeneAmp® PCR

System 9700 with the following four steps: 1) 1 cycle at 94°C for 1 min; 2) 5 cycles of 94°C for 30 sec and 65°C for 2.5 min; 3) 5 cycles of 94°C for 30 sec and 60°C for 2.5 min; and 4) 25 cycles of 94°C for 30 sec and 58°C for 2.5 min. The PCR fragments were isolated using gel electrophoresis and a QIAquick® kit as directed. The PCR fragments were directly cloned into the pCR®2.1-TOPO® vector, as directed. Clones 18, 19, 20, 24, 25, and 26 were sequenced with the M13 primers that hybridized to the vector DNA (SEQ ID NOs:25 and 26) and primers designed to anneal to the cDNA sequence (SEQ ID NOs:112, 120, 121, and 122).

5-RSA-2	GGAAAGAGTAATTGATCAGAGCCATC	(SEQ ID NO:120)
5-RSA-4	CGCCGAAGCCTCTCGCCTCACATTTCC	(SEQ ID NO:121)
3-RSA-4	GGAAATGTGAGGCGAGAGGCTTCGGCG	(SEQ ID NO:122)

The polynucleotide sequences of clones 18, 19, 20, 24, 25, and 26 are set out in SEQ ID NOs:123-128, respectively.

Clones 18, 19, 20, 24, 25, 26 and clone *Rsa*I were aligned using the Sequencher™ program. The polynucleotide sequence of the cDNA clones confirmed that there were no introns present in the *Rsa*I clone sequence. Base pairs 1-596 of clones 18, 19, 20, 24, 25, and 26 were compiled into a consensus nucleotide sequence with bp 59-596 of clone *Rsa*I that is designated 5'-RSA/cDNA and is set out in SEQ ID NO:129. The polynucleotide sequence of 5'-RSA/cDNA does not include nucleotide sequence 3' to base pair 597 of clones 18, 19, 20, 24, 25, 26, which is discussed below. The polynucleotide sequence of 5'-RSA/cDNA also does not include bp 1-58 of clone *Rsa*I, as this nucleotide sequence was not confirmed in the cDNA clone sequence. In the consensus nucleotide sequence of 5'-RSA/cDNA, every base pair was present at the corresponding position in 6 of the 7 clones, except nucleotide position 47 in which the consensus base pair was present at the corresponding position in 4 of the 7 clones.

The alignment of clones 18, 19, 20, 24, 25, and 26 identified a difference in the nucleotide sequence 3' to base pair 597 (reference position in SEQ ID NOs:123-128). All of the aligned clones contain one copy of a 10 base pair sequence (GAGCTGGCAG; SEQ ID NO:130) located at nt 588-597 (SEQ ID NOs:123-128). Clones 19 and 26 have a second copy of the sequence GAGCTGGCAG repeated

directly adjacent to the first copy (nt 598-607) (SEQ ID NOs:124 and 128). Clone *Rsa*I, clone *Tru*9I, and clone *Bfa*I also have two copies of the sequence GAGCTGGCAG directly adjacent to each other (nt 646-665 in clone *Rsa*I, (SEQ ID NO:115); nt 227-246 in clone *Tru*9I (SEQ ID NO:113); and nt 80-99 in clone *Bfa*I (SEQ ID NO:117)). Clones 18, 20, 24, and 25 do not have the second copy of the sequence GAGCTGGCAG. The presence or absence of the second copy of the sequence GAGCTGGCAG could result from an error in PCR amplification caused by *Taq* polymerase. Direct sequencing of genomic DNA can be used to verify this prediction. The presence or absence of the second copy of the sequence GAGCTGGCAG could also be caused by replication and/or repair proteins present in the bacteria used to propagate the cloned DNA. Direct sequencing of PCR products can be used to verify this prediction. The presence or absence of the second copy of the sequence GAGCTGGCAG could also result from alternative 3'-splice acceptor usage. This possibility seems unlikely since the sequences surrounding the GAGCTGGCAG sequence do not show high resemblance to the consensus sequence for exon/intron/exon borders [Lewin, *supra*]. In addition, clones generated from PCR amplification of genomic DNA have been isolated that contain only one copy of the GAGCTGGCAG sequence (Genomic 1X; SEQ ID NO:131) as well as clones containing two copies of the GAGCTGGCAG sequence (clones *Rsa*I (SEQ ID NO:115) *Tru*9I (SEQ ID NO:113) and *Bfa*I (SEQ ID NO:117)). The presence or absence of the second copy of the sequence GAGCTGGCAG may also be a polymorphism present in the human population. In this case, expression of a long and short form of the TANK2 protein would be possible, as discussed below.

The presence of two copies of the sequence GAGCTGGCAG produces a long form of the TANK2 protein. Clones 19, 26, *Rsa*I, *Tru*9I, and *Bfa*I were aligned with 5'-RSA/cDNA and 2B.1/2D.1/30B.2A/5'-RACE using the Sequencher™ program. A consensus polynucleotide sequence designated tankyrase2-long was developed from the alignment and is set out in SEQ ID NO:132. The sequence of tankyrase2-long was determined to have an ORF from nt 103-4386, with the first methionine beginning at nt 229. An in-frame stop codon (beginning at nt 100) was present upstream of the putative initiating methionine. Assuming that this residue is the

initiating methionine, the ORF of tankyrase2-long encodes a protein of 1385 amino acids (designated TANK2-LONG; SEQ ID NO:133) with a predicted molecular weight of 149,892 Da.

The presence of one copy of the sequence GAGCTGGCAG produces a short form of the TANK2 protein. Clones 18, 20, 24, and 25 were aligned with 5'-RSA/cDNA and 2B.1/2D.1/30B.2A/5'-RACE using the Sequencher™ program. A consensus polynucleotide sequence designated tankyrase2-short was developed from the alignment and is set out in SEQ ID NO:134. The sequence of tankyrase2-short was determined to have an ORF from nt 513-4376, with the first methionine beginning at nt 876. An in frame stop codon (beginning at nt 510) was present upstream of the putative initiating methionine. Assuming this residue to be the initiating methionine, the ORF of tankyrase2-short encoded a 1166 amino acid protein (designated TANK2-SHORT; SEQ ID NO:135) with a predicted molecular weight of 126,908 Da. TANK2-SHORT is 219 amino acids shorter at the amino terminal end than TANK2-LONG. The putative initiating methionine of TANK2-SHORT corresponds to a methionine at position 120 of TANK2-LONG. Excluding the first 219 amino acids of TANK2-LONG, TANK2-LONG and TANK2-SHORT are identical.

The tankyrase1 gene (SEQ ID NO:3) encodes a protein TANK1 (SEQ ID NO:4) containing a carboxyl-terminal catalytic domain that has homology to the catalytic domain of human PARP1. The polynucleotide sequence of parp1 is set out in SEQ ID NO:136, and the amino acid sequence of PARP1 is set out in SEQ ID NO:137. The catalytic domain of TANK1 (aa 1176-1314 of SEQ ID NO:4) is homologous to the catalytic domain of PARP1 (aa 854-1014 of SEQ ID NO:137) and contains PARP catalytic activity (Smith et al., *supra*). Similarly, the putative catalytic domain of TANK2-LONG (aa 1242-1382 of SEQ ID NO:133) and TANK2-SHORT (aa 1023-1161 of SEQ ID NO:135) is highly homologous to the catalytic domain of TANK1 (130 of 139 amino acids are the same; 94% identity).

The central domain of TANK1 contains 24 ankyrin repeats, indicating that TANK1 might belong to the ankyrin family of proteins that bridge integral membrane proteins to the cytoskeleton [Bennett, *J Biol Chem* 267: 8703-6 (1992)]. The ankyrin

repeat domain of TANK1 (aa 181-1110 of SEQ ID NO:4) is significantly homologous to a central domain of TANK2-LONG (aa 242-1078 of SEQ ID NO:133) and TANK2-SHORT (aa 23-859 of SEQ ID NO:135) (692 of 837 amino acids are the same; 83% identity).

Within the ankyrin repeat domain of TANK1 is a binding site for the telomeric repeat binding factor-1 (TRF1) (Smith et al., *supra*) that functions to regulate the length of telomeres [van Steensel and de Lange, *Nature* 385:740-3 (1997)]. The TRF1 binding domain of TANK1 (aa 436-797 of SEQ ID NO:4) is significantly homologous to a region of TANK2-LONG (aa 497-858 of SEQ ID NO:133) and TANK2-SHORT (aa 278-639 of SEQ ID NO:135) (297 of 364 amino acids are the same; 82% identity).

TANK1 also contains a sterile alpha module (SAM) domain [Smith et al., *supra*] that is thought to be involved in protein-protein interactions [Ponting, *Protein Sci* 4: 1928-30 (1995); Schultz et al., *Protein Sci* 6: 249-53 (1997)]. A region of TANK2-LONG (aa 1089-1154 of SEQ ID NO:133) and TANK2-SHORT (aa 870-935 of SEQ ID NO:135) is homologous to the SAM domain of TANK1 (aa 1023-1088 of SEQ ID NO:4) (50 of 66 amino acids are the same; 76% identity).

A comparison of several putative functional domains of TANK2 (catalytic domain, ankyrin repeats, TRF-1 binding domain, and SAM domain) with TANK1 is discussed above. The additional amino terminal sequence contained in TANK2-LONG (all residues amino terminal to the ankyrin repeats, i.e., aa 1-241 of SEQ ID NO:133) allows for a comparison with the amino terminus of TANK1. The amino terminus of TANK1 contains homopolymeric runs of histidines, prolines, and serines (HPS domain, i.e., aa 1-180 of SEQ ID NO:4) [Smith et al., *supra*]. The amino terminus of TANK2-LONG does not contain a HPS domain nor is it significantly homologous with the amino terminus of TANK1. The amino terminus of TANK2-LONG is also 61 amino acid residues longer than TANK1 and is composed of 48.1% non-polar residues, 32.4% polar residues, and 19.5% charged residues.

TANK2-SHORT is 219 amino acid residues shorter than TANK2-LONG and only contains 22 amino acid residues amino terminal to the ankyrin repeats. Interestingly, the *Drosophila melanogaster* tankyrase gene (GenBank® Accession No.

AF132196; SEQ ID NO:138) encodes a putative protein designated dTANK (SEQ ID NO:139) that only contains 21 amino acid residues amino terminal to its ankyrin repeats. The amino terminal ends of TANK2-SHORT and dTANK are not significantly homologous, although the two proteins do share homology in the other putative functional domains discussed above. The catalytic domain of TANK2-SHORT (aa 1023-1161 of SEQ ID NO:135) is homologous to a region of dTANK (aa 1033-1171 of SEQ ID NO:139) (113 of 139 amino acids are the same; 81% identity). The putative ankyrin repeat domain of TANK2-SHORT (aa 23-859 of SEQ ID NO:135) is significantly homologous to a central domain of dTANK (aa 22-875 SEQ ID NO:139) (545 of 858 amino acids are the same; 64% identity). The putative TRF1 binding domain of TANK2-SHORT (aa 278-639 of SEQ ID NO:135) is significantly homologous to a region of dTANK (aa 277-633 SEQ ID NO:139) (241 of 364 amino acids are the same; 66% identity). The putative SAM domain of TANK2-SHORT (aa 870-935 of SEQ ID NO:135) is significantly homologous to a region of dTANK (aa 886-951 of SEQ ID NO:139) (31 of 66 amino acids are the same; 66% identity).

EXAMPLE 3

Preparation of Antibodies Immunoreactive with TANK2 Polypeptides

The present invention provides for antibodies with specificity for TANK2 polypeptides. Antibodies to TANK2 may be produced by any method known in the art typically including, for example, the immunization of laboratory animals with preparations of purified native TANK2, purified recombinant TANK2, purified recombinant fragments of TANK2, or synthetic peptides derived from the TANK2 predicted amino acid sequence. To maximize the probability of obtaining antibodies with appropriate specificity for TANK2, regions of the polypeptide may be selected for use as an immunogen based upon differences in those regions between TANK1 and TANK2. For example, alignment of TANK1 and TANK2 demonstrates that a region consisting of aa 969-974 of TANK1 (SEQ ID NO:4) is substantially different from the corresponding region (aa 1030-1042) of TANK2-LONG (SEQ ID NO:133). In addition, the amino terminal domains of TANK1 (aa 1-180 of SEQ ID NO:4) and TANK2-LONG (aa 1-241 of SEQ ID NO:133) are substantially different, as discussed

above. These regions can be expressed as truncated polypeptides in an appropriate expression system for use as immunogen or to test polyclonal or monoclonal antibody preparations. Similar approaches can be applied to other regions of the TANK2 polypeptide. Likewise, synthetic peptides can be made to correspond to various regions of differences and such peptides can be utilized to generate specific polyclonal or monoclonal antibodies by methods known in the art. For examples, see discussions in Harlow et al. (1988), *supra*.

Alignment of TANK1 and TANK2 indicated that a region of TANK2-LONG consisting of aa 1030-1042 (SEQ ID NO:133) was substantially different than the corresponding region of TANK1 (aa 969-974 of SEQ ID NO:4). A peptide, designated ICEC #2, having this TANK2 sequence, was synthesized by AnaSpec Inc. (San Jose, CA) for use as an immunogen in antibody development. Peptide ICEC #2 was conjugated to KLH using Imject® Maleimide Activated Carrier Proteins (Pierce, #77106) following the manufacturer's protocol.

Each of four 6 to 12 week old Balb/c mice were pre-bled on day 0 and immunized by subcutaneous injection of 50 μg per mouse of KLH-ICEC-2 peptide in Freund's complete adjuvant. Subsequent boosts were made on day 21 and 42 in Freund's incomplete adjuvant. Mice were test bled on day 52 and the bleeds were screened by ELISA, using standard methods, on plates coated with KLH-ICEC-2 peptide. Specific antibody was detected using goat anti-mouse IgG(fc) horseradish peroxidase (HRP) conjugate. Mouse #3616 was given pre-fusion boosts on day 118 and 119 with 50 μg KLH-ICEC-2 peptide in PBS. The spleen was removed and fused on day 122.

Splenocytes were fused to NS-1 cells in a ratio of 5:1 by standard methods using polyethylene glycol 1500 (Boehringer Mannheim/Roche Molecular Biochemicals) [Harlow et al. (1988), *supra*]. The fused cells were resuspended in 250 mL RPMI containing 15% FBS, 100 mM sodium hypoxanthine, 0.4 mM aminopterin, 16 mM thymidine (HAT) (Gibco BRL, Rockville, MD), 10 units/mL IL-6 (Boehringer Mannheim/Roche Molecular Biochemicals) and 1.5 X 10^6 murine thymocytes/mL. The suspension was dispensed into twelve and a half 96-well flat bottom tissue culture plates (Corning, United Kingdom) at 200 μL/well. Cells in plates were fed on

days 4, 5, and 6 post fusion by aspirating approximately 100 µL from each well and adding 100 µL/well plating medium described above except lacking thymocytes.

Supernatants from the fused cells were screened on day 7-12, initially by ELISA on the immunogen, as described above. To ensure clonality, positive wells chosen from the fusion were subcloned 3 times by limiting dilution, using media lacking aminopterin. Cloning was completed for one fusion, 345C, which remained reactive to the immunizing protein. Isotyping of the antibody was performed by standard ELISA methods, using goat anti-mouse IgG1, IgG2a, IgG2b, and IgG3 HRP conjugates as detecting antibodies. The clone 345C was IgG1.

Western analysis was also used to test immunoreactivity of 345C to TANK2. 1 X 10^7 non-proliferating human PBL cells were pelleted by centrifugation and lysed by addition of 0.5 mL Buffer D [0.1% NP 40, 0.1% TX-100, 100 mM KCl, 20 mM HEPES, pH 7.9, 0.2 mM EDTA, 0.2 mM EGTA, 1.0 mM dithiothreitol (DTT), and protease inhibitor cocktail tablets, (Boehringer Mannheim/Roche Molecular Biochemicals)]. Lysates were sonicated (Sonifier® 250, Branson Ultrasonics Corp., Danbury, CT) at 20% output for 30 seconds and clarified in a 4°C microfuge for 5 min and the pellets discarded. Mouse IgG (2.5 µg) or 0.5 mL 345C mAb culture supernatant was added to the lysates and they were incubated for 90 min at 4°C. Immune complexes were collected by precipitation with 30 µL protein G-Agarose slurry (Pierce) with gentle rocking for 30 minutes at 4°C. Pellets were washed 4X in Buffer D, resuspended in 25 µL 1X SDS Sample buffer [50 mM Tris-HCl, pH 6.8, 2% SDS, 0.1% bromophenol blue, 10% glycerol, and 100 mM DDT], and heated for 5 min at 100°C.

Samples were electrophoresed on 8% Tris-Glycine polyacrylamide gels (Novex, San Diego, CA) at 60 mA for 30 min, as described by the manufacturer. Gels were transferred to Immobilon-P transfer membrane (Millipore, Bedford, MA) using a Bio-Rad (Hercules, CA) semi-dry blotting apparatus at 150 mA for 90 min as described by the manufacturer. Blots were then blocked in TBST buffer (Tris buffered saline, pH 7.5 and 0.5% Tween®) containing 5.0% nonfat dry milk for 20-30 min at room temperature. Primary mAb 345C culture supernatant was then added at a 1:2 dilution to TBST containing 1.0% nonfat dry milk and blots were incubated at

room temperature for 90 min. Following 4 washes with TBST, secondary antibody (goat anti-mouse IgG HRP conjugate, Bio-Rad) was added at a 1/3,000 dilution in TBST containing 1.0% nonfat dry milk and blots were incubated for 30 min at room temperature. Blots were again washed 4X in TBST followed by incubation in ECL detection reagents (Amersham Life Sciences, Uppsala, Sweden) as described by the manufacturer, followed by exposure to X-ray film. Positive signals of approximately the expected size for TANK2-LONG and TANK2-SHORT were obtained. The entire procedure is repeated to obtain more strongly immunoreactive monoclonal antibodies.

EXAMPLE 4

Analysis of Tank2 Expression by Northern Blot Hybridization

In order to identify cell and tissue types that express tankyrase2 mRNA, Northern blot analysis was performed using commercially prepared multi-tissue Northern blots (Clontech). The DNA probe template was amplified by PCR using a primer (5-Tank2-15; SEQ ID NO:140) corresponding to the sense strand of FB2B.1 polynucleotide sequence (nt 2330-2349 of SEQ ID NO:88) and a primer (3-Tank2-18; SEQ ID NO:141) corresponding to the antisense strand of FB2B.1 polynucleotide sequence (nt 2656-2675 of SEQ ID NO:88).

5-Tank2-15	GGCCTGAAGGTATGGTCGAT	(SEQ ID NO:140)
3-Tank2-18	TGAGGGCATTACAGTTTGTT	(SEQ ID NO:141)

The PCR reaction contained 100 ng FB2B.1 cDNA, 0.25 μM each primer, 0.20 mM dNTPs, 1X PCR buffer, and 1 μL of Clontech Advantage® polymerase mix. The reactions were performed in a GeneAmp® PCR System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec; and 3) 1 cycle at 72°C for 7 min. The PCR fragment (designated Tank2-Nprobe; SEQ ID NO:142) was isolated using gel electrophoresis and a QIAquick® kit as directed. Tank2-Nprobe was labeled with ^{32}P with a Random Primed DNA Labeling Kit (Boehringer Mannheim/Roche Molecular Biochemicals) as directed and used to probe Clontech multi-tissue Northern blots. Prehybridization with Clontech's ExpressHyb™ DNA Hybridization solution was performed at 68°C for 30 min. Hybridization with labeled probe was performed for 1 hr at 68°C in

ExpressHyb™. The blots were washed three times at room temperature in buffer containing 2X SSC and 0.05% SDS and then washed two times at 50°C in buffer containing 0.1X SSC and 0.1% SDS prior to autoradiography.

The tissue Northern blot contained an approximately 6.3 kb band whose signal was strongest in placenta, PBL, ovary, and spleen and was present in pancreas, kidney, skeletal muscle, liver, lung, brain, heart, colon, small intestine, testis, prostate, and thymus.

EXAMPLE 5

Analysis of Tank2 Expression by *in situ* Hybridization

Expression of tankyrase2 was examined in tissue sections by *in situ* hybridization as described below.

Preparation of probes

A probe for tankyrase2 *in situ* hybridization was generated using procedures routinely practiced in the art. A primer (5-Tank2-15p; SEQ ID NO:143) corresponding to the sense strand of FB2B.1 polynucleotide sequence (nt 2330-2349 of SEQ ID NO:88) and a primer (3-Tank2-18p; SEQ ID NO:144) corresponding to the antisense strand of FB2B.1 polynucleotide sequence (nt 2656-2675 of SEQ ID NO:88) were synthesized for use in a PCR reaction using FB2B.1 as the template.

5-Tank2-15p GCCGAATTCGGCCTGAAGGTATGGTCGAT (SEQ ID NO:143)

3-Tank2-18p GCCGAATTCTAGATGAGGGCATTACAGTTTGTT (SEQ ID NO:144)

The PCR reaction contained 100 ng FB2B.1 cDNA, 0.5 μM each primer, 0.25 mM dNTPs, 1X PCR buffer, and 2.5 U of *PfuTurbo*® polymerase mix (Stratagene). The reactions were performed in a GeneAmp® PCR System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 25 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 1 min; and 3) 1 cycle at 72°C for 7 min. The PCR fragment was digested with *Eco*RI, isolated using gel electrophoresis and a QIAquick® kit, and subcloned into a Bluescript® vector (Stratagene). The clone, designated Tank2-

ISprobe, was sequenced with the M13 primers designed to anneal to the vector (SEQ ID NOs:25 and 26) and the sequence is set out in SEQ ID NO:145. Tank2-ISprobe was digested with *Xho*I and transcribed (see below) with T3 polymerase to generate an antisense probe. A sense probe was generated by digesting Tank2-ISprobe with *Bam*HI and transcribing with T7 polymerase.

To compare the tissue expression of tankyrase2 with tankyrase1, a tankyrase1 probe was generated. The tankyrase1 probe corresponds to a region in the 3' untranslated sequence of the tankyrase1 gene. The 3' untranslated sequence of tankyrase1, designated 3-Tank1UT, is set out in SEQ ID NO:146. A primer (5-Tank1-7p; SEQ ID NO:147) corresponding to the sense strand of 3-Tank1UT polynucleotide sequence (nt 407-426 of SEQ ID NO:146) and a primer (3-Tank1-13p; SEQ ID NO:148) corresponding to the antisense strand of 3-Tank1UT polynucleotide sequence (nt 742-767 of SEQ ID NO:146) were synthesized for use in a PCR reaction using 3-Tank1UT as the template.

5-Tank1-7p GCCGAATTCCTTGTTTTGATTTGCCAGA (SEQ ID NO:147)

3-Tank1-13p GCCGAATTCCGGCTTTGACTTCTCTGAATTTAGG (SEQ ID NO:148)

The PCR reaction contained 100 ng 3-Tank1UT cDNA, 0.5 μM each primer, 0.25 mM dNTPs, 1X PCR buffer, and 2.5 U of *PfuTurbo*® polymerase mix (Stratagene). The reactions were performed in a GeneAmp® PCR System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 1 min; and 3) 1 cycle at 72°C for 7 min. The PCR fragment was digested with *Eco*RI, isolated using gel electrophoresis and a QIAquick® kit, and subcloned into a Bluescript® vector (Stratagene). The clone, designated Tank1-ISprobe, was sequenced with the M13 primers (SEQ ID NOs:25 and 26) and the sequence is set out in SEQ ID NO:149. Tank1-ISprobe was digested with *Bam*HI and transcribed with T7 polymerase to generate an antisense probe. A sense probe was generated by digesting Tank1-ISprobe with *Xho*I and transcribing with T3 polymerase.

The Tank1-IS probe and Tank2-ISprobe were transcribed using a RNA Transcription kit (Stratagene) in a reaction containing 5 μL of 5X transcription buffer,

30 mM DTT, 0.8mM each ATP, CTP, GTP, 40 U RNase Block II, 12.5 U T3 or T7 polymerase, 300 ng linearized plasmid template, and 50 μCi ^{35}S-UTP (greater than 1000 Ci/mmol, Amersham, Arlington Heights, IL). The mixture was incubated at 37°C for 1 hr, after which the template DNA was removed by addition of 1 μL of RNase-free DNase I (Stratagene) and incubated for 15 min at 37°C. A Quick Spin G50 RNA column (5' → 3' Inc., Boulder, CO) was prepared according to the manufacturer's suggested protocol. Twenty-five microliters (25 μL) of dH_2O was added to the probe and it was placed in the center of the column and the column centrifuged for 4 min at 1100 rpm in a desk top centrifuge. The column flow-through was mixed with 50 μL dH_2O, 2 μL of a 10mg/mL tRNA solution, 10 μL 3 M sodium acetate, and 200 μL 100% ethanol (VWR, So. Plainfield, NJ) and the resulting mixture was incubated at -20°C overnight. The probe solution was centrifuged for 15 min at 4°C, the supernatant was removed, and the pellet was resuspended in 40 μL 1X TBE [90 mM Tris-Borate and 2 mM EDTA (pH 8.0)] containing 1 μL of 0.1 M DTT. The probe was stored at -70°C until the *in situ* hybridization was performed.

Preparation of tissue samples and *in situ* hybridization

Tissues (National Disease Research Interchange, Philadelphia, PA and Cooperative Human Tissue Network, Philadelphia, PA) were sectioned at 6 μm and placed on Superfrost® Plus slides (VWR). Sections were fixed for 20 min at 4°C in 4% paraformaldehyde (Sigma, St. Louis, MO). The slides were rinsed in three changes of 1X CMF-PBS, dehydrated with three successive washes with 70% ethanol, 95% ethanol, and 100% ethanol, and dried for 30 min at room temperature. The slides were placed in 70% formamide (J.T. Baker, Phillpsburg, NJ) in 2X SSC for 2 min at 70°C, rinsed in 2X SSC at 4°C, dehydrated through 70%, 95%, and 100% ethanol washes, and dried for 30 min at room temperature. Slides were placed in an airtight box containing a piece of filter paper saturated with box buffer containing 50% formamide in 4X SSC. The probes, as described above, were individually prepared by mixing 4 X 10^5 cpm/ tissue section with 5 μL of a 10 mg/mL tRNA solution per section and heating the mixture at 95°C for 3 min. Ice-cold rHB2 buffer [10% dextran sulfate (Sigma), 50% formamide, 100 mM DTT (Boehringer

Mannheim/Roche Molecular Biochemicals), 0.3 M NaCl (Sigma), 20 mM Tris, pH 7.5, 5 mM EDTA (Sigma), and 1X Denhardt's solution (Sigma)] was added to the probe mixture to bring the final volume to 60 μL/section. The probe solution was then added to the tissue sections. The slides were incubated at 50°C for 12-16 hr. Following hybridization, the slides were washed once in 4X SSC containing 10 mM DTT for 1 hr at room temperature, once in 50% deionized formamide, 1X SSC, and 1 mM DTT for 40 min at 60°C, once in 2X SSC for 30 min at room temperature, and once in 0.1X SSC for 30 min at room temperature. The sections were dehydrated through 70%, 95%, and 100% ethanol washes and air dried for 30 min. The slides were dipped in Kodak (Rochester, NY) NTB2 nuclear emulsion at 45°C for 3 hr at room temperature in the dark and stored in the dark at 4°C with desiccant until time of development.

The slides were rinsed in dH_2O and stained with hematoxylin and eosin by transfer of the slides through a series of the following steps: 5 min in formaldehyde/alcohol (100 mL formaldehyde, 900 mL 80% ethanol); three rinses in water for a total of 2 min; 5 min in 0.75% Harris hematoxylin (Sigma); three rinses in water for a total of 2 min; one dip in 1% HCl/50% ethanol; one rinse in water; four dips in 1% lithium carbonate; 10 min in tap water; 2 min in 0.5% eosin (Sigma); three rinses in water for a total of 2 min; 2 min in 70% ethanol; three 1 min rinses in 95% ethanol; two 1 min rinses in 100% ethanol; and two 2 min rinses in xylene. Slides were mounted with cytoseal 60 (Stephens Scientific, Riverdale, NJ).

The signals obtained with the antisense tankyrase1 or antisense tankyrase2 probes were compared to the control signals obtained by the respective sense probes and any signal specific to the antisense tankyrase1 or antisense tankyrase2 probe was assumed to represent tankyrase1 or tankyrase2 expression, respectively. Both tankyrase1 and tanyrase2 signal was detected in most areas of the human testis, including the spermatogonia and spermatocytes. Tankyrase1 signal was detected in the red pulp of the human spleen while tankyrase2 signal was detected in the white pulp of the human spleen. The probes for tankyrase1 and tankyrase2 are used to detect expression in other tissues in a similar manner. Tankyrase1 signal was detected uniformly in mouse embryo, with the highest signal present in the skin. Tankyrase2

signal was also detected uniformly in mouse embryo, with the highest signal present in the mesenchymal areas and in the brain.

EXAMPLE 6

Identification of a Tankyrase2 Binding Partner

As described above, TANK1 interacts with the telomere-specific DNA binding protein TRF1 [Smith et al., (1998), *supra*]. The polynucleotide sequence of TRF1 is set out in SEQ ID NO:150, and the amino acid sequence of TRF1 is set out in SEQ ID NO:151. The yeast two-hybrid system [Hollenburg et al., *Mol Cell Biol* 15:3813-22 (1995)] was used to determine if TANK2 also interacts with TRF1. In this yeast two-hybrid system, the yeast strain L40 has been engineered to contain multiple LexA binding sites upstream of the *HIS3* and beta-galactosidase genes. Interaction of one protein fused to LexA (created in the BTM116 vector) with a second protein fused to the VP16 activation domain (created in the VP16 vector) results in the expression of *HIS3*, allowing yeast growth in media lacking histidine. Interaction of the two proteins also results in the expression of the beta-galactosidase gene, which can be measured in a colorometric assay [Breeden and Nasmyth, *Cold Spring Harbor Symp Quant Biol* 643-650 (1985)]

The TANK1 binding domain of TRF1, here designated TRF1-TankBD, has been mapped to an amino terminal region of TRF1. TRF1-TankBD was amplified by PCR using a primer (5-TRF1; SEQ ID NO:152) corresponding to the sense strand of TRF1 polynucleotide sequence (nt 1-24 of SEQ ID NO:150) and a primer (3-TRF1; SEQ ID NO:153) corresponding to the antisense strand of TRF1 polynucleotide sequence (nt 184-201 of SEQ ID NO:150).

5-TRF1 GCCCCGGGGATCCTCATGGCGGAGGATGTTTCCTCAGCG (SEQ ID NO:152)

3-TRF1 TCCCGGGGATCCTCACACCAGGCCCGCGTCCTC (SEQ ID NO:153)

The PCR reaction contained 5 μL Clontech human testis Marathon®-Ready cDNA, 0.20 μM each primer, 0.20 mM dNTPs, 1X PCR buffer, and 1 μL of Clontech Advantage® polymerase mix. The reactions were performed in a GeneAmp® PCR

System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec; and 3) 1 cycle at 72°C for 7 min. The PCR fragment was digested with *BamHI*, isolated using gel electrophoresis and a QIAquick® kit as directed, and subcloned into the BTM116 vector. TRF1-TankBD was sequenced with the M13 reverse primer designed to anneal to the vector (SEQ ID NO:26) and a primer designed to anneal to the cDNA sequence (SEQ ID NO:153). The polynucleotide sequence of TRF1-TankBD is set out in SEQ ID NO:154 and the amino acid sequence is set out in SEQ ID NO:155.

As described above, the TRF1 binding domain of TANK1 is very homologous to a region of TANK2 comprised of aa 497-858 of SEQ ID NO:133. The polynucleotide region corresponding to this domain of TANK2, designated Tank2-TRF1BD, was amplified in a PCR reaction with a primer (5-T2/TRF1BD; SEQ ID NO:156) corresponding to the sense strand of the tank2 polynucleotide sequence (nt 1717-1742 of SEQ ID NO:132) and a primer (3-T2/TRF1BD; SEQ ID NO:157) corresponding to the antisense strand of the tank2 polynucleotide sequence (nt 2765-2805 of SEQ ID NO:132).

5-T2/TRF1BD CGCAGGATCCCCTTCACTCCTCTTCATGAGGCAGCTTC (SEQ ID NO:156)

3-T2/TRF1BD GGATCCGCTAAATATCTGTATCTCCATCTTTAACAA GATCCAAAGGAG (SEQ ID NO:157)

The PCR reaction contained 5 µL Clontech human testis Marathon®-Ready cDNA, 0.5 µM each primer, 0.25 mM dNTPs, 1X PCR buffer, and 2.5 U of *PfuTurbo*® polymerase mix (Stratagene). The reactions were performed in a GeneAmp® PCR System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 55°C for 2 min, and 72°C for 2 min; and 3) 1 cycle at 72°C for 7 min. The PCR fragment was isolated using gel electrophoresis and a QIAquick® kit as directed, and subcloned into the pCR-BluntII™-TOPO® vector (Invitrogen). Tank2-TRF1BD was digested from the pCR-BluntII™-TOPO® with *Bam*HI, and subcloned into the VP16 vector. The Tank2-TRF1BD clone was sequenced with primers designed to adhere to the vector sequence: M13 forward (SEQ ID NO:25) and 009 (SEQ ID NO:158).

009	GCCGACTTCGAGTTTGAGCAG	(SEQ ID NO:158)

The polynucleotide sequence is set out in SEQ ID NO:159 and the amino acid sequence is set out in SEQ ID NO:160.

Co-transformation of L40 with the TRF1-TankBD and Tank2-TRF1BD plasmids indicated that like TANK1, TANK2 binds to TRF1.

EXAMPLE 7

Measurement of TANK2 Biological Activity

Construction of Expression Plasmids

The primary structure of the tankyrase2 polypeptide suggests that TANK2, like TANK1, will have poly(ADP-ribose) polymerase activity. The PARP activity of TANK2, or some substructure thereof, can be measured by the ability of that component to incorporate the ADP-ribose unit from NAD into polymers of ADP-ribose coupled to a protein substrate. For example, TANK1 adds polymers of ADP-ribose to the TRF-1 protein in molecular assays [Smith et al., *supra*]. TANK2 is expected to also perform this function and/or to ADP-ribosylate another substrate or substrates. The demonstration of such activity on a given substrate is readily accomplished by the skilled artisan [see, for example, Smith et al., *supra*].

Structural differences in TANK1 and TANK2 suggest the possibility that TANK2 may have different protein substrate specificity than does TANK1. As demonstrated by the observation that TANK1 binds to TRF-1 and poly ADP-ribosylates TRF-1, it is anticipated that protein substrates of TANK2 can be identified by their ability to bind to TANK2. Additional substrates that bind TANK2 can be identified by a number of methods as described elsewhere in this application.

A fusion protein, designated PARP1A/TANK2B, containing aa 1-662 of PARP1 (SEQ ID NO:137) fused upstream of aa 996-1385 of TANK2 (SEQ ID NO:133) was used in the measurement of TANK2 poly(ADP-ribose) polymerase activity. PARP1A/TANK2B contained the DNA binding domain (aa 1-373 of SEQ ID NO:137) and automodification domain (aa 373-525 of SEQ ID NO:137) of PARP1 and the putative catalytic domain of TANK2 (aa 1242-1382 of SEQ ID NO:133).

The PARP1A piece of the fusion protein was amplified by PCR using a primer (Sal-PARP1; SEQ ID NO:161) corresponding to the sense strand of parp1 polynucleotide sequence (nt 1-30 of SEQ ID NO:136) and a primer (revMlu-PARP1; SEQ ID NO:162) corresponding to the antisense strand of parp1 polynucleotide sequence (nt 1957-1985 of SEQ ID NO:136).

Sal-PARP1 CGTCGACCCATGGCGGAGTCTTCGGATAAGCTCTATCGA

(SEQ ID NO:161)

revMlu-PARP1

GGAAACGCGTTTGGTGCCAGGATTTACTGTCAGCTTCTT

(SEQ ID NO:162)

The PCR reaction contained 0.5 μL of human thymus and testis QUICK-Clone™ cDNA (Clontech), 0.25 μM each primer, 0.20 mM dNTPs, 1X PCR buffer, and 1 μL of Clontech Advantage® polymerase mix. The reactions were performed in a GeneAmp® (PE Applied Biosystems) with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 60°C for 2 min, and 72°C for 2 min; and 3) 1 cycle at 72°C for 7 min. The PCR fragment (designated parp1A) was isolated using gel electrophoresis and a QIAquick® kit as directed. Parp1A was subcloned into the pTrcHis2™-TOPO® vector (Invitrogen) as directed. Parp1A was digested from pTrcHis2™-TOPO® with *Sal*I and *Mlu*I, the fragment isolated using gel electrophoresis and a QIAquick® kit, and saved for further subcloning described below.

The TANK2B piece of the fusion protein was amplified by PCR using a primer (forMlu-TANK2; SEQ ID NO:163) corresponding to the sense strand of tank2 polynucleotide sequence (nt 3214-3240 of SEQ ID NO:132) and a primer (TANK2-Strep-Not; SEQ ID NO:164) corresponding to the antisense strand of tank2 polynucleotide sequence (nt 4350-4383 of SEQ ID NO:132).

ForMlu-TANK2

CTTAAACGCGTTGAAGGACAAACACCTTTAGATTTAGTT

(SEQ ID NO:163)

TANK2-Strep-Not
GTCGAAAGCGGCCGCTTAGCCTCCGAACTGTGGATGCC
TCCACGCTCCATCGACCATACCTTCAGGCCTCATAATCTGG
(SEQ ID NO:164)

The PCR reaction contained 100 ng 2B.1 cDNA, 0.25 μM each primer, 0.20 mM dNTPs, 1X PCR buffer, and 1 μL of Clontech Advantage® polymerase mix. The reactions were performed in a GeneAmp® PCR System 9700 with the following steps: 1) 1 cycle at 94°C for 1 min; 2) 30 cycles of 94°C for 30 sec, 60°C for 2 min, and 72°C for 2 min; and 3) 1 cycle at 72°C for 7 min. The PCR fragment (designated tank2B) was isolated using gel electrophoresis and a QIAquick® kit as directed. Tank2B was subcloned into the pCDNA3.1/NT-GFP-TOPO® vector (Invitrogen) as directed. Tank2B was digested from pCDNA3.1/NT-GFP-TOPO® with *Mlu*I and *Not*I and subcloned with *Sal*I/*Mlu*I digested parp1A (see above) into a pFASTBAC vector (Gibco BRL), which had previously been digested with *Sal*I and *Not*I. The resultant plasmid was designated pFB-PARP1A/TANK2B.

pFB-PARP1A/TANK2B was sequenced with primers designed to anneal to the vector sequence (SEQ ID NOs:165-166) and primers designed to anneal to the cDNA sequence (SEQ ID NOs:55, 60, and 66, *supra*, and SEQ ID NOs:167-176).

Vector Primers

FastBac for	TTTGTTCGCCCAGACTC	(SEQ ID NO:165)
FastBac rev	TATGTTTCAGGTTCAGGGGGAG	(SEQ ID NO:166)

cDNA Primers

P1	GCGGAAGCTGGAGGAGTGAC	(SEQ ID NO:167)
P2	GTCACTCCTCCAGCTTCCGC	(SEQ ID NO:168)
P3	AAGCCCTGAAGAAGCAGCTC	(SEQ ID NO:169)
P4	GAGCTGCTTCTTCAGGGCTT	(SEQ ID NO:170)
P5	CAGACACCCAACCGGAAGGA	(SEQ ID NO:171)
P6	TCCTTCCGGTTGGGTGTCTG	(SEQ ID NO:172)
P7	TCCGCCTCCACCAAGAGCCT	(SEQ ID NO:173)
P8	AGGCTCTTGGTGGAGGCGGA	(SEQ ID NO:174)
P9	TGGCCTGGTGGACATCGTTA	(SEQ ID NO:175)
P10	TAACGATGTCCACCAGGCCA	(SEQ ID NO:176)

The nucleotide sequence of PARP1A/TANK2B is set out in SEQ ID NO:177 and the amino acid sequence of PARP1A/TANK2B is set out in SEQ ID NO:178. PARP1A/TANK2B consists of the following regions: a HIS tag leader region at aa 1-36; a PARP1 region at aa 37-698; a spacer region at aa 699-700; a TANK2 region at aa 701-1090; and a *Strep*-tag region at aa 1091-1099.

Production of Recombinant Viral Stocks and Protein Purification

PARP1A/TANK2B recombinant viral stock was produced using the FastBac system (Gibco BRL) according to the manufacturer's suggested protocol and protein expression was carried out as follows. *Sf9* cells were grown at 27°C in CCM3 medium (Hyclone, Logan, UT) containing 50 U/mL penicillin and 50 µg/mL streptomycin sulfate (Gibco BRL). Exponentially growing cells were infected at a multiplicity of infection of approximately 0.5 virus per cell and incubated for 48 hr. Cells were collected by centrifugation at 1000 X *g* for 15 min, and the pellets were frozen and stored at -80°C until use.

For protein purification, reagents were obtained from Sigma unless otherwise indicated. Cells were lysed in Lysis buffer [25 mM Tris-HCl, pH 9.0, 50 mM glucose, 10 mM EDTA, 1 mM 2-mercaptoethanol, 1 mM PMSF, 100 µM antipain, and 2 µg/mL aprotinin] by sonication. Igepal CA-630 (final concentration of 0.2%), Tween®-20 (final concentration of 0.2%), and NaCl (final concentration of 0.5 M) were added to the Lysis buffer and the samples were agitated for 30 min at 4°C. The supernatants were collected after centrifugation at 20,000 X *g* for 20 min at 4°C, at which time they were treated with 1 mg/mL protamine sulfate and allowed to stir for 1 hr at 4°C. The supernatants were collected after centrifugation at 4,000 X *g* for 20 min at 4°C at which time the protein was precipitated with 70% ammonium sulfate. Protein pellets were collected by centrifugation at 20,000 X *g* for 15 min at 4°C and resuspended in Re-suspension buffer [100 mM Tris-HCl, pH 7.4, 0.5 mM EDTA, 10% glycerol, 1 mM PMSF, and 12 mM 2-mercaptoethanol].

Proteins were first purified via the HIS tag using Talon® Superflow metal affinity resin (Clontech) and eluted with 200 mM imidazole (Clontech) as directed. The protein elutions were next purified using a 3-aminobenzamide Affi-Gel® matrix

(Bio-Rad Laboratories) prepared as described elsewhere [D'Amours et al., *Anal Biochem* 249:106-8 (1997)]. Proteins were eluted with 10 mM 3-methoxybenzamide in Elution buffer [50 mM Tris-HCl, pH 7.5, 0.3 M NaCl, 10 mM 2-mercaptoethanol, 1 mM PMSF, 100 μM antipain, and 2 μg/mL aprotinin]. The proteins were dialyzed 4 X in 1 L Dialysis buffer [50 mM Tris-HCl, pH 8.0, 1 mM DTT, 4 mM $MgCl_2$, 10 mM EDTA, 1 mM PMSF, and 2 μg/mL aprotinin). Glycerol was added to a final concentration of 10% and the proteins were stored at -80°C.

Poly(ADP-ribose) polymerase activity

For poly(ADP-ribose) polymerase activity assays, reagents were obtained from Sigma unless otherwise indicated. PARP1A/TANK2B (250 ng) protein was incubated for 10 min at room temperature in Assay buffer (total volume of 20 μL) [100 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 10% glycerol, 1.5 mM DTT (Boehringer Mannheim/Roche Molecular Biochemicals), 2.5 μM unlabeled NAD^+, 16.7 μg/mL *E. coli* Strain B DNA, and 0.33 μCi γ-[^{32}P]-NAD^+ (NEN, Boston, MA). Reactions were stopped by boiling in SDS running buffer and separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Autoradiography was used to visualize labeled protein. Addition of poly(ADP-ribose) polymers to protein substrate results in an increase in molecular weight of the protein, and consequently causes the protein to run higher on SDS PAGE. Also, the level of poly(ADP-ribose) polymers added to the protein substrate can vary with each single protein molecule, resulting in labeled proteins with different molecular weights, which appears on the autoradiography film as a ladder or smear [for example, see Smith et al. *Science* 282:2484-7 (1998)]. PARP1A/TANK2B possessed intrinsic poly(ADP-ribose) polymerase activity as shown by its ability produce poly(ADP-ribose) polymers. The PARP1A/TANK2B poly(ADP-ribose) polymerase reaction produced a ladder of labeled protein from approximately 136 kDa to 250 kDa.

All publications and patent documents cited in this specification are incorporated herein by reference for all that they disclose.

While the present invention has been described with specific reference to certain preferred embodiments for purposes of clarity and understanding, it will be apparent to the skilled artisan that further changes and modifications may be practiced within the scope of the invention as it is defined in the claims set forth below. Accordingly, no limitations should be placed on the invention other than those specifically recited in the claims.

WHAT IS CLAIMED IS:

1. A purified and isolated tankyrase2 polypeptide.

2. The polypeptide according to Claim 1, comprising the amino acid sequence defined in SEQ ID NO:133.

3. The polypeptide according to Claim 1, comprising the amino acid sequence defined in SEQ ID NO:135.

4. A polynucleotide encoding the polypeptide according to Claim 1.

5. The polynucleotide according to Claim 4, comprising the coding region of the nucleotide sequence defined in SEQ ID NO:132.

6. The polynucleotide according to Claim 4, comprising the coding region of the nucleotide sequence defined in SEQ ID NO:134.

7. A polynucleotide selected from the group consisting of:

(a) the polynucleotide according to Claim 4,

(b) a polynucleotide complementary to the polynucleotide of (a), and

(c) a polynucleotide that hybridizes under moderately stringent hybridization conditions to the polynucleotide of (a) or (b).

8. The polynucleotide according to Claim 7, wherein the polynucleotide is a DNA molecule or an RNA molecule.

9. The polynucleotide according to Claim 8, further comprising a detectable label moiety.

10. An expression construct, comprising the polynucleotide according to Claim 4.

11. A host cell transformed or transfected with the expression construct according to Claim 10.

12. The polynucleotide according to Claim 4, wherein the polynucleotide is operatively linked to a heterologous promoter.

13. A host cell, comprising the polynucleotide according to Claim 12.

14. A method for producing a tankyrase2 polypeptide, comprising the steps of:

a) growing the host cell according to Claim 11 or 13 under conditions appropriate for expression of the polypeptide; and

b) isolating the polypeptide from the host cell or the medium in which the host cell is grown.

15. An antibody that is specifically immunoreactive with the polypeptide according to Claim 1.

16. The antibody according to Claim 15, wherein the antibody is selected from the group consisting of monoclonal antibodies, polyclonal antibodies, single chain antibodies (scFv antibodies), chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, CDR-grafted antibodies, Fab fragments, Fab' fragments, $F(ab')_2$ fragments, and Fv fragments.

17. A cell line that produces an antibody according to Claim 15.

18. An anti-idiotype antibody that is specifically immunoreactive with an antibody according to Claim 15.

19. A method for identifying a binding partner of a tankyrase2 polypeptide, comprising:

a) contacting the tankyrase2 polypeptide with a test compound under conditions that permit binding of the tankyrase2 polypeptide and the test compound;

b) detecting binding of the test compound and the tankyrase2 polypeptide; and

c) identifying the test compound as a binding partner of the tankyrase2 polypeptide.

20. The method according to Claim 19, wherein said specific binding partner selectively or specifically modulates a biological activity of the tankyrase2 polypeptide.

21. A method for identifying a specific binding partner of a tankyrase2 polynucleotide, comprising:

a) contacting the tankyrase2 polynucleotide with a test compound under conditions that permit binding of the tankyrase2 polynucleotide and the test compound;

b) detecting binding of the test compound and the tankyrase2 polynucleotide; and

c) identifying the test compound as a specific binding partner of the tankyrase2 polynucleotide.

22. The method according to Claim 21, wherein said binding partner selectively or specifically modulates activity of the tankyrase2 polynucleotide.

23. A method of treating an animal having a medical condition mediated by poly(ADP-ribose) polymerase activity, comprising administering to said animal a tankyrase2 inhibitory compound in an amount effective for inhibiting tankyrase2 activity in said animal.

24. The method according to Claim 23, wherein said medical condition is associated with growth of neoplastic tissue.

25. The method according to Claim 24, wherein said neoplastic tissue is a cancer selected from the group consisting of carcinomas, sarcomas, leukemias, and lymphomas.

26. The method according to Claim 25, wherein said cancer is selected from the group consisting of ACTH-producing tumor, acute lymphocytic leukemia, acute nonlymphocytic leukemia, cancer of the adrenal cortex, bladder cancer, brain cancer, breast cancer, cervical cancer, chronic lymphocytic leukemia, chronic myelocytic leukemia, colorectal cancer, cutaneous T-cell lymphoma, endometrial cancer, esophageal cancer, Ewing's sarcoma, gallbladder cancer, hairy cell leukemia, head and neck cancer, Hodgkin's lymphoma, Kaposi's sarcoma, kidney cancer, liver cancer, lung cancer (small and non-small cell), malignant peritoneal effusion, malignant pleural effusion, melanoma, mesothelioma, multiple myeloma, neuroblastoma, glioma, non-Hodgkin's lymphoma, osteosarcoma, ovarian cancer, ovarian (germ cell) cancer, pancreatic cancer, penile cancer, prostate cancer, retinoblastoma, skin cancer, soft tissue sarcoma, squamous cell carcinomas, stomach cancer, testicular cancer, thyroid cancer, trophoblastic neoplasms, uterine cancer, vaginal cancer, cancer of the vulva, and Wilm's tumor.

SEQUENCE LISTING

<110> Christenson, Erik
DeMaggio, Anthony J
Goldman, Phyllis S
McElligott, David L

<120> Tankyrase2 Materials and Methods

<130> 36559

<140>
<141>

<150> 60/141,582
<151> 1999-06-29

<160> 178

<170> PatentIn Ver. 2.1

<210> 1
<211> 3507
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (2)..(3508)

<400> 1
g gcc agg atc atg tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc 49
Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
1 5 10 15

gcg agc gcc gcg gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc 97
Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
20 25 30

gag gcg tgc cgc aac ggg gac gtg gaa cga gtc aag agg ctg gtg acg 145
Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
35 40 45

cct gag aag gtg aac agc cgc gac acg gcg ggc agg aaa tcc acc ccg 193
Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
50 55 60

ctg cac ttc gcc gca ggt ttt ggg cgg aaa gac gta gtt gaa tat ttg 241
Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
65 70 75 80

ctt cag aat ggt gca aat gtc caa gca cgt gat gat ggg ggc ctt att 289
Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
85 90 95

cct ctt cat aat gca tgc tct ttt ggt cat gct gaa gta gtc aat ctc 337
Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
100 105 110

ctt ttg cga cat ggt gca gac ccc aat gct cga gat aat tgg aat tat 385
Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
115 120 125

act cct ctc cat gaa gct gca att aaa gga aag att gat gtt tgc att 433
Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
130 135 140

```
gtg ctg tta cag cat gga gct gag cca acc atc cga aat aca gat gga     481
Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145                 150                 155                 160

agg aca gca ttg gat tta gca gat cca tct gcc aaa gca gtg ctt act     529
Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
                165                 170                 175

ggt gaa tat aag aaa gat gaa ctc tta gaa agt gcc agg agt ggc aat     577
Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
            180                 185                 190

gaa gaa aaa atg atg gct cta ctc aca cca tta aat gtc aac tgc cac     625
Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
        195                 200                 205

gca agt gat ggc aga aag tca act cca tta cat ttg gca gca gga tat     673
Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
    210                 215                 220

aac aga gta aag att gta cag ctg tta ctg caa cat gga gct gat gtc     721
Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225                 230                 235                 240

cat gct aaa gat aaa ggt gat ctg gta cca tta cac aat gcc tgt tct     769
His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
                245                 250                 255

tat ggt cat tat gaa gta act gaa ctt ttg gtc aag cat ggt gcc tgt     817
Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
            260                 265                 270

gta aat gca atg gac ttg tgg caa ttc act cct ctt cat gag gca gct     865
Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
        275                 280                 285

tct aag aac agg gtt gaa gta tgt tct ctt ctc tta agt tat ggt gca     913
Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
    290                 295                 300

gac cca aca ctg ctc aat tgt cac aat aaa agt gct ata gac ttg gct     961
Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305                 310                 315                 320

ccc aca cca cag tta aaa gaa aga tta gca tat gaa ttt aaa ggc cac    1009
Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
                325                 330                 335

tcg ttg ctg caa gct gca cga gaa gct gat gtt act cga atc aaa aaa    1057
Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
            340                 345                 350

cat ctc tct ctg gaa atg gtg aat ttc aag cat cct caa aca cat gaa    1105
His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
        355                 360                 365

aca gca ttg cat tgt gct gct gca tct cca tat ccc aaa aga aag caa    1153
Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
    370                 375                 380

ata tgt gaa ctg ttg cta aga aaa gga gca aac atc aat gaa aag act    1201
Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385                 390                 395                 400

aaa gaa ttc ttg act cct ctg cac gtg gca tct gag aaa gct cat aat    1249
Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
                405                 410                 415
```

gat gtt gtt gaa gta gtg gtg aaa cat gaa gca aag gtt aat gct ctg 1297
Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
420 425 430

gat aat ctt ggt cag act tct cta cac aga gct gca tat tgt ggt cat 1345
Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
435 440 445

cta caa acc tgc cgc cta ctc ctg agc tat ggg tgt gat cct aac att 1393
Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
450 455 460

ata tcc ctt cag ggc ttt act gct tta cag atg gga aat gaa aat gta 1441
Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465 470 475 480

cag caa ctc ctc caa gag ggt atc tca tta ggt aat tca gag gca gac 1489
Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
485 490 495

aga caa ttg ctg gaa gct gca aag gct gga gat gtc gaa act gta aaa 1537
Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
500 505 510

aaa ctg tgt act gtt cag agt gtc aac tgc aga gac att gaa ggg cgt 1585
Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg
515 520 525

cag tct aca cca ctt cat ttt gca gct ggg tat aac aga gtg tcc gtg 1633
Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val
530 535 540

gtg gaa tat ctg cta cag cat gga gct gat gtg cat gct aaa gat aaa 1681
Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
545 550 555 560

gga ggc ctt gta cct ttg cac aat gca tgt tct tat gga cat tat gaa 1729
Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
565 570 575

gtt gca gaa ctt ctt gtt aaa cat gga gca gta gtt aat gta gct gat 1777
Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp
580 585 590

tta tgg aaa ttt aca cct tta cat gaa gca gca gca aaa gga aaa tat 1825
Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr
595 600 605

gaa att tgc aaa ctt ctg ctc cag cat ggt gca gac cct aca aaa aaa 1873
Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys
610 615 620

aac agg gat gga aat act cct ttg gat ctt gtt aaa gat gga gat aca 1921
Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr
625 630 635 640

gat att caa gat ctg ctt agg gga gat gca gct ttg cta gat gct gcc 1969
Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala
645 650 655

aag aag ggt tgt tta gcc aga gtg aag aag ttg tct tct cct gat aat 2017
Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn
660 665 670

gta aat tgc cgc gat acc caa ggc aga cat tca aca cct tta cat tta 2065
Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu
675 680 685

```
gca gct ggt tat aat aat tta gaa gtt gca gag tat ttg tta caa cac     2113
Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His
    690                 695                 700

gga gct gat gtg aat gcc caa gac aaa gga gga ctt att cct tta cat     2161
Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His
705                 710                 715                 720

aat gca gca tct tac ggg cat gta gat gta gca gct cta cta ata aag     2209
Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys
                725                 730                 735

tat aat gca tgt gtc aat gcc acg gac aaa tgg gct ttc aca cct ttg     2257
Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu
            740                 745                 750

cac gaa gca gcc caa aag gga cga aca cag ctt tgt gct ttg ttg cta     2305
His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu
        755                 760                 765

gcc cat gga gct gac ccg act ctt aaa aat cag gaa gga caa aca cct     2353
Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro
    770                 775                 780

tta gat tta gtt tca gca gat gat gtc agc gct ctt ctg aca gca gcc     2401
Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala
785                 790                 795                 800

atg ccc cca tct gct ctg ccc tct tgt tac aag cct caa gtg ctc aat     2449
Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn
                805                 810                 815

ggt gtg aga agc cca gga gcc act gca gat gct ctc tct tca ggt cca     2497
Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro
            820                 825                 830

tct agc cca tca agc ctt tct gca gcc agc agt ctt gac aac tta tct     2545
Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser
        835                 840                 845

ggg agt ttt tca gaa ctg tct tca gta gtt agt tca agt gga aca gag     2593
Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu
    850                 855                 860

ggt gct tcc agt ttg gag aaa aag gag gtt cca gga gta gat ttt agc     2641
Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser
865                 870                 875                 880

ata act caa ttc gta agg aat ctt gga ctt gag cac cta atg gat ata     2689
Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile
                885                 890                 895

ttt gag aga gaa cag atc act ttg gat gta tta gtt gag atg ggg cac     2737
Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His
            900                 905                 910

aag gag ctg aag gag att gga atc aat gct tat gga cat agg cac aaa     2785
Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys
        915                 920                 925

cta att aaa gga gtc gag aga ctt atc tcc gga caa caa ggt ctt aac     2833
Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn
    930                 935                 940

cca tat tta act ttg aac acc tct ggt agt gga aca att ctt ata gat     2881
Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp
945                 950                 955                 960
```

ctg tct cct gat gat aaa gag ttt cag tct gtg gag gaa gag atg caa 2929
Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln
965 970 975

agt aca gtt cga gag cac aga gat gga ggt cat gca ggt gga atc ttc 2977
Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe
980 985 990

aac aga tac aat att ctc aag att cag aag gtt tgt aac aag aaa cta 3025
Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu
995 1000 1005

tgg gaa aga tac act cac cgg aga aaa gaa gtt tct gaa gaa aac cac 3073
Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His
1010 1015 1020

aac cat gcc aat gaa cga atg cta ttt cat ggg tct cct ttt gtg aat 3121
Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn
1025 1030 1035 1040

gca att atc cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt ggt 3169
Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly
1045 1050 1055

atg ttt gga gct ggc att tat ttt gct gaa aac tct tcc aaa agc aat 3217
Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn
1060 1065 1070

caa tat gta tat gga att gga gga ggt act ggg tgt cca gtt cac aaa 3265
Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys
1075 1080 1085

gac aga tct tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg gta 3313
Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val
1090 1095 1100

acc ttg gga aag tct ttc ctg cag ttc agt gca atg aaa atg gca cat 3361
Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His
1105 1110 1115 1120

tct cct cca ggt cat cac tca gtc act ggt agg ccc agt gta aat ggc 3409
Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly
1125 1130 1135

cta gca tta gct gaa tat gtt att tac aga gga gaa cag gct tat cct 3457
Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro
1140 1145 1150

gag tat tta att act tac cag att atg agg cct gaa ggt atg gtc gat 3505
Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp
1155 1160 1165

gga 3508
Gly

<210> 2
<211> 1169
<212> PRT
<213> Homo sapiens

<400> 2
Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
1 5 10 15

Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
20 25 30

Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
35 40 45

Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
50 55 60

Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
65 70 75 80

Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
85 90 95

Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
100 105 110

Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
115 120 125

Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
130 135 140

Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145 150 155 160

Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
165 170 175

Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
180 185 190

Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
195 200 205

Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
210 215 220

Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225 230 235 240

His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
245 250 255

Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
260 265 270

Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
275 280 285

Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
290 295 300

Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305 310 315 320

Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
325 330 335

Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
340 345 350

His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
355 360 365

Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
370 375 380

Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385 390 395 400

Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
405 410 415

Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
420 425 430

Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
435 440 445

Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
450 455 460

Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465 470 475 480

Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
485 490 495

Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
500 505 510

Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg
515 520 525

Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val
530 535 540

Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
545 550 555 560

Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
565 570 575

Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp
580 585 590

Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr
595 600 605

Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys
610 615 620

Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr
625 630 635 640

Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala
645 650 655

Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn
660 665 670

Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu
675 680 685

Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His
690 695 700

Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His
705 710 715 720

Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys
725 730 735

Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu
740 745 750

His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu
755 760 765

Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro
770 775 780

Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala
785 790 795 800

Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn
805 810 815

Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro
820 825 830

Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser
835 840 845

Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu
850 855 860

Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser
865 870 875 880

Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile
885 890 895

Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His
900 905 910

Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys
915 920 925

Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn
930 935 940

Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp
945 950 955 960

Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln
965 970 975

Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe
980 985 990

Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu
995 1000 1005

Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His
1010 1015 1020

Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn
1025 1030 1035 1040

Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly
1045 1050 1055

Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn
1060 1065 1070

Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys
1075 1080 1085

Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val
1090 1095 1100

Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His
1105 1110 1115 1120

Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly
1125 1130 1135

Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro
1140 1145 1150

Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp
1155 1160 1165

Gly

<210> 3
<211> 3984
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(3981)

<400> 3
atg gcg gcg tcg cgt cgc tct cag cat cat cac cac cat cat caa caa 48
Met Ala Ala Ser Arg Arg Ser Gln His His His His His His Gln Gln
1 5 10 15

cag ctc cag ccc gcc cca ggg gct tca gcg ccg ccg ccg cca cct cct 96
Gln Leu Gln Pro Ala Pro Gly Ala Ser Ala Pro Pro Pro Pro Pro Pro
20 25 30

ccc cca ctc agc cct ggc ctg gcc ccg ggg acc acc cca gcc tct ccc 144
Pro Pro Leu Ser Pro Gly Leu Ala Pro Gly Thr Thr Pro Ala Ser Pro
35 40 45

acg gcc agc ggc ctg gcc ccc ttc gcc tcc ccg cgg cac ggc cta gcg 192
Thr Ala Ser Gly Leu Ala Pro Phe Ala Ser Pro Arg His Gly Leu Ala
50 55 60

ctg ccg gag ggg gat ggc agt cgg gat ccg ccc gac agg ccc cga tcc 240
Leu Pro Glu Gly Asp Gly Ser Arg Asp Pro Pro Asp Arg Pro Arg Ser
65 70 75 80

ccg gac ccg gtt gac ggt acc agc tgt tgc agt acc acc agc aca atc 288
Pro Asp Pro Val Asp Gly Thr Ser Cys Cys Ser Thr Thr Ser Thr Ile
85 90 95

tgt acc gtc gcc gcc gct ccc gtg gtc cca gcg gtt tct act tca tct 336
Cys Thr Val Ala Ala Ala Pro Val Val Pro Ala Val Ser Thr Ser Ser
100 105 110

gcc gct ggg gtc gct ccc aac cca gcc ggc agt ggc agt aac aat tca 384
Ala Ala Gly Val Ala Pro Asn Pro Ala Gly Ser Gly Ser Asn Asn Ser
115 120 125

ccg tcg tcc tct tct tcc ccg act tct tcc tca tct tcc tct cca tcc 432
Pro Ser Ser Ser Ser Ser Pro Thr Ser Ser Ser Ser Ser Ser Pro Ser
130 135 140

tcc cct gga tcg agc ttg gcg gag agc ccc gag gcg gcc gga gtt agc 480
Ser Pro Gly Ser Ser Leu Ala Glu Ser Pro Glu Ala Ala Gly Val Ser
145 150 155 160

agc aca gca cca ctg ggg cct ggg gca gca gga cct ggg aca ggg gtc 528
Ser Thr Ala Pro Leu Gly Pro Gly Ala Ala Gly Pro Gly Thr Gly Val
165 170 175

cca gca gtg agc ggg gcc cta cgg gaa ctg ctg gag gcc tgt cgc aat 576
Pro Ala Val Ser Gly Ala Leu Arg Glu Leu Leu Glu Ala Cys Arg Asn
180 185 190

ggg gac gtg tcc cgg gta aag agg ctg gtg gac gcg gca aac gta aat 624
Gly Asp Val Ser Arg Val Lys Arg Leu Val Asp Ala Ala Asn Val Asn
195 200 205

```
gca aag gac atg gcc ggc cgg aag tct tct ccc ctg cac ttc gct gca    672
Ala Lys Asp Met Ala Gly Arg Lys Ser Ser Pro Leu His Phe Ala Ala
    210                 215                 220

ggt ttt gga agg aag gat gtt gta gaa cac tta cta cag atg ggt gct    720
Gly Phe Gly Arg Lys Asp Val Val Glu His Leu Leu Gln Met Gly Ala
225                 230                 235                 240

aat gtc cac gct cgt gat gat gga ggt ctc atc ccg ctt cat aat gcc    768
Asn Val His Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His Asn Ala
                245                 250                 255

tgt tct ttt ggc cat gct gag gtt gtg agt ctg tta ttg tgc caa gga    816
Cys Ser Phe Gly His Ala Glu Val Val Ser Leu Leu Leu Cys Gln Gly
            260                 265                 270

gct gat cca aat gcc agg gat aac tgg aac tat aca cct ctg cat gaa    864
Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu His Glu
        275                 280                 285

gct gct att aaa ggg aag atc gat gtg tgc att gtg ctg ctg cag cac    912
Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu Gln His
    290                 295                 300

gga gct gac cca aac att cgg aac act gat ggg aaa tca gcc ctg gac    960
Gly Ala Asp Pro Asn Ile Arg Asn Thr Asp Gly Lys Ser Ala Leu Asp
305                 310                 315                 320

ctg gca gat cct tca gca aaa gct gtc ctt aca ggt gaa tac aag aaa    1008
Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr Lys Lys
                325                 330                 335

gac gaa ctc cta gaa gct gct agg agt ggt aat gaa gaa aaa cta atg    1056
Asp Glu Leu Leu Glu Ala Ala Arg Ser Gly Asn Glu Glu Lys Leu Met
            340                 345                 350

gct tta ctg act cct cta aat gtg aat tgc cat gca agt gat ggg cga    1104
Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly Arg
        355                 360                 365

aag tcg act cct tta cat cta gca gcg ggc tac aac aga gtt cga ata    1152
Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val Arg Ile
    370                 375                 380

gtt cag ctt ctt ctt cag cat ggt gct gat gtt cat gca aaa gac aaa    1200
Val Gln Leu Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
385                 390                 395                 400

ggt gga ctt gtg cct ctt cat aat gca tgt tca tat gga cat tat gaa    1248
Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
                405                 410                 415

gtc aca gaa ctg cta cta aag cat gga gct tgt gtt aat gcc atg gat    1296
Val Thr Glu Leu Leu Leu Lys His Gly Ala Cys Val Asn Ala Met Asp
            420                 425                 430

ctc tgg cag ttt act cca ctg cac gag gct gct tcc aag aac cgt gta    1344
Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val
        435                 440                 445

gaa gtc tgc tct ttg tta ctt agc cat ggc gct gat cct acg tta gtc    1392
Glu Val Cys Ser Leu Leu Leu Ser His Gly Ala Asp Pro Thr Leu Val
    450                 455                 460

aac tgc cat ggc aaa agt gct gtg gat atg gct cca act ccg gag ctt    1440
Asn Cys His Gly Lys Ser Ala Val Asp Met Ala Pro Thr Pro Glu Leu
465                 470                 475                 480
```

```
agg gag aga ttg act tat gaa ttt aaa ggt cat tct tta cta caa gca      1488
Arg Glu Arg Leu Thr Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala
                485                 490                 495

gcc aga gaa gca gac tta gct aaa gtt aaa aaa aca ctc gct ctg gaa      1536
Ala Arg Glu Ala Asp Leu Ala Lys Val Lys Lys Thr Leu Ala Leu Glu
            500                 505                 510

atc att aat ttc aaa caa ccg cag tct cat gaa aca gca ctg cac tgt      1584
Ile Ile Asn Phe Lys Gln Pro Gln Ser His Glu Thr Ala Leu His Cys
        515                 520                 525

gct gtg gcc tct ctg cat ccc aaa cgt aaa caa gtg aca gaa ttg tta      1632
Ala Val Ala Ser Leu His Pro Lys Arg Lys Gln Val Thr Glu Leu Leu
    530                 535                 540

ctt aga aaa gga gca aat gtt aat gaa aaa aat aaa gat ttc atg act      1680
Leu Arg Lys Gly Ala Asn Val Asn Glu Lys Asn Lys Asp Phe Met Thr
545                 550                 555                 560

ccc ctg cat gtt gca gcc gaa aga gcc cat aat gat gtc atg gaa gtt      1728
Pro Leu His Val Ala Ala Glu Arg Ala His Asn Asp Val Met Glu Val
                565                 570                 575

ctg cat aag cat ggc gcc aag atg aat gca ctg gac acc ctt ggt cag      1776
Leu His Lys His Gly Ala Lys Met Asn Ala Leu Asp Thr Leu Gly Gln
            580                 585                 590

act gct ttg cat aga gcc gcc cta gca ggc cac ctg cag acc tgc cgc      1824
Thr Ala Leu His Arg Ala Ala Leu Ala Gly His Leu Gln Thr Cys Arg
        595                 600                 605

ctc ctg ctg agt tac ggc tct gac ccc tcc atc atc tcc tta caa ggc      1872
Leu Leu Leu Ser Tyr Gly Ser Asp Pro Ser Ile Ile Ser Leu Gln Gly
    610                 615                 620

ttc aca gca gca cag atg ggc aat gaa gca gtg cag cag att ctg agt      1920
Phe Thr Ala Ala Gln Met Gly Asn Glu Ala Val Gln Gln Ile Leu Ser
625                 630                 635                 640

gag agt aca cct ata cgt act tct gat gtt gat tat cga ctc tta gag      1968
Glu Ser Thr Pro Ile Arg Thr Ser Asp Val Asp Tyr Arg Leu Leu Glu
                645                 650                 655

gca tct aaa gct gga gac ttg gaa act gtg aag caa ctt tgc agc tct      2016
Ala Ser Lys Ala Gly Asp Leu Glu Thr Val Lys Gln Leu Cys Ser Ser
            660                 665                 670

caa aat gtg aat tgt aga gac tta gag ggc cgg cat tcc acg ccc tta      2064
Gln Asn Val Asn Cys Arg Asp Leu Glu Gly Arg His Ser Thr Pro Leu
        675                 680                 685

cac ttc gca gca ggc tac aac cgc gtg tct gtt gta gag tac ctg cta      2112
His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu
    690                 695                 700

cac cac ggt gcc gat gtc cat gcc aaa gac aag ggt ggc ttg gtg ccc      2160
His His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro
705                 710                 715                 720

ctt cat aat gcc tgt tca tat gga cac tat gag gtg gct gag ctt tta      2208
Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu
                725                 730                 735

gta agg cat ggg gct tct gtc aat gtg gcg gac tta tgg aaa ttt acc      2256
Val Arg His Gly Ala Ser Val Asn Val Ala Asp Leu Trp Lys Phe Thr
            740                 745                 750
```

```
cct ctc cat gaa gca gca gct aaa gga aag tat gaa atc tgc aag ctc     2304
Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu
        755                 760                 765

ctt tta aaa cat gga gca gat cca act aaa aag aac aga gat gga aat     2352
Leu Leu Lys His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn
    770                 775                 780

aca cct ttg gat ttg gta aag gaa gga gac aca gat att cag gac tta     2400
Thr Pro Leu Asp Leu Val Lys Glu Gly Asp Thr Asp Ile Gln Asp Leu
785                 790                 795                 800

ctg aaa ggg gat gct gct ttg ttg gat gct gcc aag aag ggc tgc ctg     2448
Leu Lys Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly Cys Leu
            805                 810                 815

gca aga gtg cag aag ctc tgt acc cca gag aat atc aac tgc aga gac     2496
Ala Arg Val Gln Lys Leu Cys Thr Pro Glu Asn Ile Asn Cys Arg Asp
            820                 825                 830

acc cag ggc aga aat tca acc cct ctg cac ctg gca gca ggc tat aat     2544
Thr Gln Gly Arg Asn Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn
        835                 840                 845

aac ctg gaa gta gct gaa tat ctt cta gag cat gga gct gat gtt aat     2592
Asn Leu Glu Val Ala Glu Tyr Leu Leu Glu His Gly Ala Asp Val Asn
    850                 855                 860

gcc cag gac aag ggt ggt tta att cct ctt cat aat gcg gca tct tat     2640
Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala Ser Tyr
865                 870                 875                 880

ggg cat gtt gac ata gcg gct tta ttg ata aaa tac aac acg tgt gta     2688
Gly His Val Asp Ile Ala Ala Leu Leu Ile Lys Tyr Asn Thr Cys Val
            885                 890                 895

aat gca aca gat aag tgg gcg ttt act ccc ctc cat gaa gca gcc cag     2736
Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala Ala Gln
            900                 905                 910

aaa gga agg acg cag ctg tgc gcc ctc ctc cta gcg cat ggt gca gac     2784
Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly Ala Asp
        915                 920                 925

ccc acc atg aag aac cag gaa ggc cag acg cct ctg gat ctg gca aca     2832
Pro Thr Met Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu Ala Thr
    930                 935                 940

gct gac gat atc aga gct ttg ctg ata gat gcc atg ccc cca gag gcc     2880
Ala Asp Asp Ile Arg Ala Leu Leu Ile Asp Ala Met Pro Pro Glu Ala
945                 950                 955                 960

tta cct acc tgt ttt aaa cct cag gct act gta gtg agt gcc tct ctg     2928
Leu Pro Thr Cys Phe Lys Pro Gln Ala Thr Val Val Ser Ala Ser Leu
            965                 970                 975

atc tca cca gca tcc acc ccc tcc tgc ctc tcg gct gcc agc agc ata     2976
Ile Ser Pro Ala Ser Thr Pro Ser Cys Leu Ser Ala Ala Ser Ser Ile
            980                 985                 990

gac aac ctc act ggc cct tta gca gag ttg gcc gta gga gga gcc tcc     3024
Asp Asn Leu Thr Gly Pro Leu Ala Glu Leu Ala Val Gly Gly Ala Ser
        995                 1000                1005

aat gca ggg gat ggc gcc gcg gga aca gaa agg aag gaa gga gaa gtt     3072
Asn Ala Gly Asp Gly Ala Ala Gly Thr Glu Arg Lys Glu Gly Glu Val
   1010                1015                1020
```

gct ggt ctt gac atg aat atc agc caa ttt cta aaa agc ctt ggc ctt 3120
Ala Gly Leu Asp Met Asn Ile Ser Gln Phe Leu Lys Ser Leu Gly Leu
1025 1030 1035 1040

gaa cac ctt cgg gat atc ttt gaa aca gaa cag att aca cta gat gtg 3168
Glu His Leu Arg Asp Ile Phe Glu Thr Glu Gln Ile Thr Leu Asp Val
1045 1050 1055

ttg gct gat atg ggt cat gaa gag ttg aaa gaa ata ggc atc aat gca 3216
Leu Ala Asp Met Gly His Glu Glu Leu Lys Glu Ile Gly Ile Asn Ala
1060 1065 1070

tat ggg cac cgc cac aaa tta atc aaa gga gta gaa aga ctc tta ggt 3264
Tyr Gly His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Leu Gly
1075 1080 1085

gga caa caa ggc acc aat cct tat ttg act ttt cac tgt gtt aat cag 3312
Gly Gln Gln Gly Thr Asn Pro Tyr Leu Thr Phe His Cys Val Asn Gln
1090 1095 1100

gga acg att ttg ctg gat ctt gct cca gaa gat aaa gaa tat cag tca 3360
Gly Thr Ile Leu Leu Asp Leu Ala Pro Glu Asp Lys Glu Tyr Gln Ser
1105 1110 1115 1120

gtg gaa gaa gag atg caa agt act att cga gaa cac aga gat ggt ggt 3408
Val Glu Glu Glu Met Gln Ser Thr Ile Arg Glu His Arg Asp Gly Gly
1125 1130 1135

aat gct ggc ggc atc ttc aac aga tac aat gtc att cga att caa aaa 3456
Asn Ala Gly Gly Ile Phe Asn Arg Tyr Asn Val Ile Arg Ile Gln Lys
1140 1145 1150

gtt gtc aac aag aag ttg agg gag cgg ttc tgc cac cga cag aag gaa 3504
Val Val Asn Lys Lys Leu Arg Glu Arg Phe Cys His Arg Gln Lys Glu
1155 1160 1165

gtg tct gag gag aat cac aac cat cac aat gag cgc atg ttg ttt cat 3552
Val Ser Glu Glu Asn His Asn His His Asn Glu Arg Met Leu Phe His
1170 1175 1180

ggt tct cct ttc att aat gcc att att cat aaa ggg ttt gat gag cga 3600
Gly Ser Pro Phe Ile Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg
1185 1190 1195 1200

cat gca tac ata gga gga atg ttt ggg gcc ggg att tat ttt gct gaa 3648
His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu
1205 1210 1215

aac tcc tca aaa agc aac caa tat gtt tat gga att gga gga gga aca 3696
Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr
1220 1225 1230

ggc tgc cct aca cac aag gac agg tca tgc tat ata tgt cac aga caa 3744
Gly Cys Pro Thr His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln
1235 1240 1245

atg ctc ttc tgt aga gtg acc ctt ggg aaa tcc ttt ctg cag ttt agc 3792
Met Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser
1250 1255 1260

acc atg aaa atg gcc cac gcg cct cca ggg cac cac tca gtc att ggt 3840
Thr Met Lys Met Ala His Ala Pro Pro Gly His His Ser Val Ile Gly
1265 1270 1275 1280

aga ccg agc gtc aat ggg ctg gca tat gct gaa tat gtc atc tac aga 3888
Arg Pro Ser Val Asn Gly Leu Ala Tyr Ala Glu Tyr Val Ile Tyr Arg
1285 1290 1295

```
gga gaa cag gca tac cca gag tat ctt atc act tac cag atc atg aag    3936
Gly Glu Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Lys
            1300                1305                1310

cca gaa gcc cct tcc cag acc gca aca gcc gca gag cag aag acc tag    3984
Pro Glu Ala Pro Ser Gln Thr Ala Thr Ala Ala Glu Gln Lys Thr
        1315                1320                1325


&lt;210&gt; 4
&lt;211&gt; 1327
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 4
Met Ala Ala Ser Arg Arg Ser Gln His His His His His His Gln Gln
  1               5                  10                  15

Gln Leu Gln Pro Ala Pro Gly Ala Ser Ala Pro Pro Pro Pro Pro Pro
            20                  25                  30

Pro Pro Leu Ser Pro Gly Leu Ala Pro Gly Thr Thr Pro Ala Ser Pro
        35                  40                  45

Thr Ala Ser Gly Leu Ala Pro Phe Ala Ser Pro Arg His Gly Leu Ala
    50                  55                  60

Leu Pro Glu Gly Asp Gly Ser Arg Asp Pro Pro Asp Arg Pro Arg Ser
65                  70                  75                  80

Pro Asp Pro Val Asp Gly Thr Ser Cys Cys Ser Thr Thr Ser Thr Ile
                85                  90                  95

Cys Thr Val Ala Ala Ala Pro Val Val Pro Ala Val Ser Thr Ser Ser
            100                 105                 110

Ala Ala Gly Val Ala Pro Asn Pro Ala Gly Ser Gly Ser Asn Asn Ser
        115                 120                 125

Pro Ser Ser Ser Ser Ser Pro Thr Ser Ser Ser Ser Ser Ser Pro Ser
    130                 135                 140

Ser Pro Gly Ser Ser Leu Ala Glu Ser Pro Glu Ala Ala Gly Val Ser
145                 150                 155                 160

Ser Thr Ala Pro Leu Gly Pro Gly Ala Ala Gly Pro Gly Thr Gly Val
                165                 170                 175

Pro Ala Val Ser Gly Ala Leu Arg Glu Leu Leu Glu Ala Cys Arg Asn
            180                 185                 190

Gly Asp Val Ser Arg Val Lys Arg Leu Val Asp Ala Ala Asn Val Asn
        195                 200                 205

Ala Lys Asp Met Ala Gly Arg Lys Ser Ser Pro Leu His Phe Ala Ala
    210                 215                 220

Gly Phe Gly Arg Lys Asp Val Val Glu His Leu Leu Gln Met Gly Ala
225                 230                 235                 240

Asn Val His Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His Asn Ala
                245                 250                 255

Cys Ser Phe Gly His Ala Glu Val Val Ser Leu Leu Leu Cys Gln Gly
            260                 265                 270

Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu His Glu
        275                 280                 285
```

```
Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu Gln His
    290                 295                 300

Gly Ala Asp Pro Asn Ile Arg Asn Thr Asp Gly Lys Ser Ala Leu Asp
305                 310                 315                 320

Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr Lys Lys
                325                 330                 335

Asp Glu Leu Leu Glu Ala Ala Arg Ser Gly Asn Glu Glu Lys Leu Met
            340                 345                 350

Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly Arg
        355                 360                 365

Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val Arg Ile
    370                 375                 380

Val Gln Leu Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
385                 390                 395                 400

Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
                405                 410                 415

Val Thr Glu Leu Leu Leu Lys His Gly Ala Cys Val Asn Ala Met Asp
            420                 425                 430

Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val
        435                 440                 445

Glu Val Cys Ser Leu Leu Leu Ser His Gly Ala Asp Pro Thr Leu Val
    450                 455                 460

Asn Cys His Gly Lys Ser Ala Val Asp Met Ala Pro Thr Pro Glu Leu
465                 470                 475                 480

Arg Glu Arg Leu Thr Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala
                485                 490                 495

Ala Arg Glu Ala Asp Leu Ala Lys Val Lys Lys Thr Leu Ala Leu Glu
            500                 505                 510

Ile Ile Asn Phe Lys Gln Pro Gln Ser His Glu Thr Ala Leu His Cys
        515                 520                 525

Ala Val Ala Ser Leu His Pro Lys Arg Lys Gln Val Thr Glu Leu Leu
    530                 535                 540

Leu Arg Lys Gly Ala Asn Val Asn Glu Lys Asn Lys Asp Phe Met Thr
545                 550                 555                 560

Pro Leu His Val Ala Ala Glu Arg Ala His Asn Asp Val Met Glu Val
                565                 570                 575

Leu His Lys His Gly Ala Lys Met Asn Ala Leu Asp Thr Leu Gly Gln
            580                 585                 590

Thr Ala Leu His Arg Ala Ala Leu Ala Gly His Leu Gln Thr Cys Arg
        595                 600                 605

Leu Leu Leu Ser Tyr Gly Ser Asp Pro Ser Ile Ile Ser Leu Gln Gly
    610                 615                 620

Phe Thr Ala Ala Gln Met Gly Asn Glu Ala Val Gln Gln Ile Leu Ser
625                 630                 635                 640

Glu Ser Thr Pro Ile Arg Thr Ser Asp Val Asp Tyr Arg Leu Leu Glu
                645                 650                 655
```

Ala Ser Lys Ala Gly Asp Leu Glu Thr Val Lys Gln Leu Cys Ser Ser
660 665 670

Gln Asn Val Asn Cys Arg Asp Leu Glu Gly Arg His Ser Thr Pro Leu
675 680 685

His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu
690 695 700

His His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro
705 710 715 720

Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu
725 730 735

Val Arg His Gly Ala Ser Val Asn Val Ala Asp Leu Trp Lys Phe Thr
740 745 750

Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu
755 760 765

Leu Leu Lys His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn
770 775 780

Thr Pro Leu Asp Leu Val Lys Glu Gly Asp Thr Asp Ile Gln Asp Leu
785 790 795 800

Leu Lys Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly Cys Leu
805 810 815

Ala Arg Val Gln Lys Leu Cys Thr Pro Glu Asn Ile Asn Cys Arg Asp
820 825 830

Thr Gln Gly Arg Asn Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn
835 840 845

Asn Leu Glu Val Ala Glu Tyr Leu Leu Glu His Gly Ala Asp Val Asn
850 855 860

Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala Ser Tyr
865 870 875 880

Gly His Val Asp Ile Ala Ala Leu Leu Ile Lys Tyr Asn Thr Cys Val
885 890 895

Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala Ala Gln
900 905 910

Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly Ala Asp
915 920 925

Pro Thr Met Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu Ala Thr
930 935 940

Ala Asp Asp Ile Arg Ala Leu Leu Ile Asp Ala Met Pro Pro Glu Ala
945 950 955 960

Leu Pro Thr Cys Phe Lys Pro Gln Ala Thr Val Val Ser Ala Ser Leu
965 970 975

Ile Ser Pro Ala Ser Thr Pro Ser Cys Leu Ser Ala Ala Ser Ser Ile
980 985 990

Asp Asn Leu Thr Gly Pro Leu Ala Glu Leu Ala Val Gly Gly Ala Ser
995 1000 1005

Asn Ala Gly Asp Gly Ala Ala Gly Thr Glu Arg Lys Glu Gly Glu Val
1010 1015 1020

Ala Gly Leu Asp Met Asn Ile Ser Gln Phe Leu Lys Ser Leu Gly Leu
1025 1030 1035 1040

Glu His Leu Arg Asp Ile Phe Glu Thr Glu Gln Ile Thr Leu Asp Val
1045 1050 1055

Leu Ala Asp Met Gly His Glu Glu Leu Lys Glu Ile Gly Ile Asn Ala
1060 1065 1070

Tyr Gly His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Leu Gly
1075 1080 1085

Gly Gln Gln Gly Thr Asn Pro Tyr Leu Thr Phe His Cys Val Asn Gln
1090 1095 1100

Gly Thr Ile Leu Leu Asp Leu Ala Pro Glu Asp Lys Glu Tyr Gln Ser
1105 1110 1115 1120

Val Glu Glu Glu Met Gln Ser Thr Ile Arg Glu His Arg Asp Gly Gly
1125 1130 1135

Asn Ala Gly Gly Ile Phe Asn Arg Tyr Asn Val Ile Arg Ile Gln Lys
1140 1145 1150

Val Val Asn Lys Lys Leu Arg Glu Arg Phe Cys His Arg Gln Lys Glu
1155 1160 1165

Val Ser Glu Glu Asn His Asn His His Asn Glu Arg Met Leu Phe His
1170 1175 1180

Gly Ser Pro Phe Ile Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg
1185 1190 1195 1200

His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu
1205 1210 1215

Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr
1220 1225 1230

Gly Cys Pro Thr His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln
1235 1240 1245

Met Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser
1250 1255 1260

Thr Met Lys Met Ala His Ala Pro Pro Gly His His Ser Val Ile Gly
1265 1270 1275 1280

Arg Pro Ser Val Asn Gly Leu Ala Tyr Ala Glu Tyr Val Ile Tyr Arg
1285 1290 1295

Gly Glu Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Lys
1300 1305 1310

Pro Glu Ala Pro Ser Gln Thr Ala Thr Ala Ala Glu Gln Lys Thr
1315 1320 1325

<210> 5
<211> 460
<212> DNA
<213> Homo sapiens

<400> 5
gaactgtctt cagtagttag ttcaagtgga acagagggtg cttccagttt ggagaaaaag 60
gaggttccag gagtagattt tagcataact caattcgtaa ggaatcttgg acttgagcac 120
ctaatggata tatttnagag agaacagatc actttggatg tattagttga gatggggcac 180
aaggagctga aggagattgg aatcaatgct tatggacata ggcacaaact aattaaagga 240
gtcgagagac ttatctccgg acaacaaggt cttaacccat atttaacttt gaacacctct 300

ggtagtggaa caattcttat agatctgtct cctgatgata aagagtttca gtctgtggag 360
gaagagatgc aaagtacagt tcgagagcac agagatggag gtcatgcagg tggaatcttc 420
aacagataca atattctcaa gattcagaag gtttgtaaca 460

<210> 6
<211> 42
<212> PRT
<213> Homo sapiens

<400> 6
Gly Thr Ile Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser
1 5 10 15

Val Glu Glu Glu Met Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly
20 25 30

His Ala Gly Gly Ile Phe Asn Arg Tyr Asn
35 40

<210> 7
<211> 564
<212> DNA
<213> Homo sapiens

<400> 7
tgctatttca tgggtctcct tttgtgaatg caattatcca caaaggcttt gatgaaaggc 60
atgcgtacat aggtggtatg tttggagctg gcatttattt tgctgaaaac tcttccaaaa 120
gcaatcaata tgtatatgga attggaggag gtactgggtg tccagttcac aaagacagat 180
cttgttacat ttgccacagg agnctgctct tttgccgggt aaccttggga aagtctttcc 240
tgcagttcag tgcaatgaaa atggcacatt ctcctccagg tcatcactca gtcactggta 300
ggcccagtgt aaatggccta gcattagctg aatatgttat ttacagagga gaacaggtaa 360
tgtagtttta tttgttcatc ttcaaaantg ctaggggagg catactttaa cttttttatta 420
atctcttgaa ttgacaagac ntttgcctta acgggntttt ttaaaatttt atttgggggt 480
atttttcagtt tgggaagtta caaatagtaa agagatttc ttattaccct tacccggntt 540
ccnaatgtta tattttgttc cctt 564

<210> 8
<211> 118
<212> PRT
<213> Homo sapiens

<400> 8
Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly Phe
1 5 10 15

Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr
20 25 30

Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly
35 40 45

Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile Cys
50 55 60

His Arg Xaa Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu
65 70 75 80

Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His Ser
85 90 95

Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr Val
100 105 110

Ile Tyr Arg Gly Glu Gln
115

<210> 9
<211> 397
<212> DNA
<213> Homo sapiens

<400> 9
aacagagtta acttgaacct tttatatgtt atgcattgat tctaacaaac tgtaatgccc 60
tcaacagaac taattttact aatacaatac tgtgttcttt aaaacacagc atttacactg 120
aatacaattt catttgtaaa actgtaaata agagcttttg tactagccca gtatttattt 180
acattgcttt gtaatataaa tctgttttag aactgcagcg gtttacaaaa ttttttcata 240
tgtattgttc atctatactt catcttacat cgtcatgatt gagtgatctt tacatttgat 300
tccagaggct atgttcagtt gttagttggg gaaagattga gttatcagat ttaatttgcc 360
gatgggagcc tttatctgtc ataggaaatc tttctca 397

<210> 10
<211> 343
<212> DNA
<213> Homo sapiens

<400> 10
cttatcctga gtatttaatt acttaccaga ttatgaggcc tgaaggtatg gtcgatggat 60
aaatagttat tttaagaaac taattccact gaacctaaaa tcatcaaagc agcagtggcc 120
tctacgtttt actcctttgc tgaaaaaaaa tcatcttgcc cacaggcctg tggcaaaagg 180
ataaaaatgt gaacgaagtt ttaacattct gacttgataa agctttaata atgtacagtg 240
ttttctaaat atttnctgtt ttttcagcac tttaacagat gccattccag ggtaaactgg 300
ggttgtctgt actaaattat aaacagggtt aactggaccc ttt 343

<210> 11
<211> 334
<212> DNA
<213> Homo sapiens

<400> 11
gcagttctaa aacagattta tattacaaag caatgtaaat aaatactggg ctagtacaaa 60
agctcttatt tacagtttta caaatgaaat tgtattcagt gtaaatgctg tgttttaaag 120
aacacagtat tgtattagta aaattagttc tgttgagggc attacagttt gttagaatca 180
atgcataaca tataaaaggt tcaagttaac tctgtttata atttagtaca gacaacccag 240
tttaacctgg aatggcatct gttaaagtgc tgaaaaaaca ggaaatattt agaaaacact 300
gtacattatt aaagctttat caagtcagaa tgtt 334

<210> 12
<211> 353
<212> DNA
<213> Homo sapiens

<400> 12
cagcaaagga gtaaaacgta gaggccactg ctgctttgat gattttaggt tcagtggaat 60
tagtttctta aaataactat ttatccatcg accatacctt caggcctcat aatctggtaa 120
gtaattaaat actcaggata agcctgttct cctctgtaaa taacatattc agctaatgct 180
aggccattta cactgggcct accagtgact gagtgatgac ctggaggaga atgtgccatt 240
ttcattgcac tgaactgcag gaaagacttt cccaaggtta cccggcaaaa gagcagctgc 300
ctgtggcaaa tgaacaagat ctgtctttgt gaactggaca cccagtacct tct 353

<210> 13
<211> 436
<212> DNA
<213> Homo sapiens

<400> 13
tttttttgc agttctaaaa cagatttata ttacaaagca atgtaaataa atactgggct 60
agtacaaaag ctcttattta cagttttaca aatgaaattg tattcagtgt aaatgctgtg 120
ttttaaagaa cacagtattg tattagtaaa attagttctg ttgagggcat tacagtttgt 180
tagaatcaat gcataacata taaaaggttc aagttaactc tgtttataat ttagtacaga 240
caacccagtt taacctggga tggcatctg ttaaagtgct ggaaaaaaca gggaaatatt 300
taggaaaaca ctggtacatt atttaaaggc tttntccaag gtcaggantg tttaaacttc 360
gtttcacatt tttatccntt tggccacggc ctgtggggcn aggatggatt ttttttccgg 420

ccaagggtgt taaacg 436

<210> 14
<211> 392
<212> DNA
<213> Homo sapiens

<400> 14
tgctatttca tgggtctcct tttgtgaatg caattatcca caaaggcttt gatgaaaggc 60
atgcgtacat aggtggtatg tttggagctg gcatttattt tgctgaaaac tcttccaaaa 120
gcaatcaata tgtatatgga attggaggag gtactgggtg tccagttcac aaagacagat 180
cttgttacat ttgccacagg cagctgctct tttgccgggt aaccttggga aagtctttcc 240
tgcagttcag tgcaatgaaa atggcacatt ctcctccagg tcatcactca gtcactggta 300
ggcccagtgt aaatggccta gcattagctg naatatgtta tttacagagg agaacaggta 360
atgtagtttt aattttgttt catcttccaa aa 392

<210> 15
<211> 317
<212> DNA
<213> Homo sapiens

<400> 15
ttttttttgc agttctaaaa cagatttata ttacaaagca atgtaaataa atactgggct 60
agtacaaaag ctcttattta cagttttaca aatgaaattg tattcagtgt aaatgctgtn 120
ttttaaagaa cacagtattg tattagtaaa attagttctg ttgagggcat tacagtttgt 180
taggaatcaa tgcataacat ataaaaggtt caagttaact ctgtttataa tttaggtaca 240
gacaacccag tttaaccggg gaatgggcat ctgttaaagt gctgaaaaaa cnggganata 300
tttaggaaaa cnctgta 317

<210> 16
<211> 485
<212> DNA
<213> Homo sapiens

<400> 16
tgcagttcta aaacagattt atattacaaa gcaatgtaaa taaatactgg gctagtacaa 60
aagctcttat ttacagtttt acaaatgaaa ttgtattcag tgtaaatgct gtgttttaaa 120
gaacacagta ttgtattagt aaaattagtt ctgttgaggg cattacagtt tgttagaatc 180
aatgcataac atataaaagg ttcaagttaa ctctgtttat aatttagtac agacaaccca 240
gtttaacctg gaatggcatc tgttaaagtg ctgaaaaaac aggaaatatt tacgaaaaca 300
ctgtacatta ttaaagcttt atcaagtcag aatgttaaac ttcgttcaca ttttatcct 360
tttgccacag gcctgtgggg caagatgatt ttttttcagc aaaggagtaa aacgtagagg 420
gccactggct gctttgatga ttttagggtt cagtgggaat tagtttccta aaataacnat 480
ttatc 485

<210> 17
<211> 291
<212> DNA
<213> Homo sapiens

<400> 17
ttncctgcag ttcagtgcaa tgaanatggc acattctcct ccaggtcatc actcagtcac 60
tggtaggccc agtgtaaatg gcctagcatt agctgaatat gttatttaca gaggagaaca 120
ggcttatcct gagtatttaa ttacttacca gattatgagg cctgaaggta tggtcgatgg 180
ataaatagtt attttaagaa actaattcca ctgaacctaa aatcatcaaa gcagcagtgg 240
cctctacgtt ttactccttt gctgaaaaaa aatcatcttg cccacaggcc t 291

<210> 18
<211> 371
<212> DNA
<213> Homo sapiens

<400> 18
cgtagaggcc actgctgctt tgatganttt tanggttcan gtggaattng tttcttaaaa 60
taactattta tccatcgacc ataccttcag gcctcataat ctggtaagta attaaatact 120

caggataagc ctgttctcct ctgtaaataa catattcagc taatgctagg ccatttacac 180
tgggcctacc agtgactgaa gtgatgcctg gggggagaat gtgccatttt cattgcactg 240
aactgcaggn aagactttcc caagggttac ccgggcaaaa gagcagctgc ctgtgggnaa 300
tgttacaagg tcttgtcttt tgtngacctn gggcaccccg taccctcctc caattccata 360
tacatatttg a 371

<210> 19
<211> 341
<212> DNA
<213> Homo sapiens

<400> 19
gaaagataca ctcaccggag aaaagaagtt tctgaagaaa accacaacca tgccaatgaa 60
cgaatgctat ttcatgggtc tcctttgtg aatgcaatta tccacaaagg ctttgatgaa 120
aggcatgcgt acataggtgg tatgtttgga gctggcattt attttgctgg aaaactcttc 180
caaaaggcaa tcaatatgta tatgggaatt gggagggagg gtactggggt gtccagtttc 240
acaaaggaca gatcttgttt acatttggcc acaggcaggc tggctctttt tgcccgggtn 300
accttggggg aagtcttttc ctggcagttt cagttgccat g 341

<210> 20
<211> 385
<212> DNA
<213> Homo sapiens

<400> 20
tactaaatta taaacagagt taacttgaac cttttatatg ttatgcattg attctaacaa 60
actgtaatgc cctcaacaga actaatttta ctaatacaat aangtgttct ttaaaacaca 120
gcatttacac tgaatacaat ttcatttgta aaactgtaaa taagagcttt tgtactagcc 180
cagtatttat ttacattgct ttgtaatata aatctgtttt aggaactgca ggcggtttac 240
aaaatttttt catatgtatt gttcatttat acttcatctt acatcgtcat ggattgaggt 300
gatctttaca tttggattcc ngggggctat ggttcaggtt gttaggttgg gggaaagggt 360
tggggtttat ccgggnttta ntttg 385

<210> 21
<211> 335
<212> DNA
<213> Homo sapiens

<400> 21
gaaggtatgg tcgatggata aatagttatt ttaagaaact aattccactg aacctaaaat 60
catcaaagca gcagtggcct ctacgtttta ctcctttgct gaaaaaaaat catcttgccc 120
acaggcctgt ggcaaaagga taaaaatgtg aacgaagttt aacattctga cttgataaag 180
ctttaataat gtacagtgtt ttctaaatat ttcctgtttt ttcagcactt taacagatgc 240
cattccgggt taaactgggg ttgtctgtac taaattatta aacagngtta acttggaacc 300
nttttatatg ttatggcctt ggttcttaac caana 335

<210> 22
<211> 388
<212> DNA
<213> Homo sapiens

<400> 22
gttttactcc tttgctgaaa aaaaatcatc ttgcccacag gcctgtggaa naaggataaa 60
aatgtgaacg aagtttaaca ttctgacttg ataaagcttt aataatgtac agtgttttct 120
aaatatttcc tgtttttca gcactttaac agatgccatt ccaggttaaa ctgggttgtc 180
tgtactaaat tataaacaga gttaacttga accttttata tgttatgcat tgattctaac 240
aaactgtaat gccctcaaca gaactaattt tactaataca atactgtgtt ctttaaaaca 300
caggcattta cactggaata caatttcatt tgttaaaact ggtaantagg agcttttgta 360
ctagcccagt atttatttac atgctttg 388

<210> 23
<211> 401
<212> DNA
<213> Homo sapiens

<400> 23
gttttactcc tttgctgaaa aaaaatcatc ttgcccacag gcctgtggaa naaggataaa 60
aatgtgaacg aagttaacat tctgacttga taaagcttta ataatgtaca gtgttttcta 120
aatatttcct gtttttcag cactttaaca gatgccattc caggttaaac tgggttgtct 180
gtactaaatt ataaacagag ttaacttgaa ccttttatat gttatgcatt gattctaaca 240
aactgtaatg ccctcaacag aactantttt acttaataca atactgtgtt ctttnaaaac 300
acaggcattt acactggaat acaattttca ttttgttaaa actggttaaa ttaaggnggc 360
tttttgtact nggccccgtn ttttatttta cattgctttg g 401

<210> 24
<211> 354
<212> DNA
<213> Homo sapiens

<400> 24
taattttact aatacaatac tgtgttcttt aaaacacagc atttacactg aatacaattt 60
catttgtaaa actgtaaata agagcttttg tactagccca gtatttattt acattgcttt 120
gtaatataaa tctgttttag aactgcagcg gtttacaaaa ttttttcata tgtattgttc 180
atctatactt catcttacat cgtcatgatt gagtgatctt tacatttgat tccagaggct 240
atgttcagtt gttagttggg aaagattgag ttatcagatt taatttgccg atgggagcct 300
ttatctgtca ttagaaatct ttctnattta agaacttatg aatatgctga agat 354

<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 25
tgtaaaacga cggccagt 18

<210> 26
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 26
ggaaacagct atgaccatg 19

<210> 27
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 27
tttgccgggt aaccttgg 18

<210> 28
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 28
ccaaggttac ccggcaaa 18

<210> 29
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 29
gtaggcccag tgtaaatg 18

<210> 30
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 30
catttacact gggcctac 18

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 31
gagtaagttg cagggcatgt 20

<210> 32
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 32
acatgccctg caacttactc 20

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 33
gaatcaccgc agttactaaa 20

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 34

tttagtaact gcggtgattc 20

<210> 35
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 35
ggcctgaagg tatggtcgat 20

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 36
atcgaccata ccttcaggcc 20

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 37
tgagggcatt acagtttgtt 20

<210> 38
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 38
taatacgaac tcactatagg g 21

<210> 39
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 39
atacactcac cggagaaa 18

<210> 40
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

```
<400> 40
tttctccggt gagtgtat                                                   18


<210> 41
<211> 1691
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(357)

<220>
<223> Description of Artificial Sequence: Sequence not
      specified as protein-coding is vector sequence

<400> 41
atg cta ttt cat ggg tct cct ttt gtg aat gca att atc cac aaa ggc      48
Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly
  1               5                  10                  15

ttt gat gaa agg cat gcg tac ata ggt ggt atg ttt gga gct ggc att      96
Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile
             20                  25                  30

tat ttt gct gaa aac tct tcc aaa agc aat caa tat gta tat gga att     144
Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile
         35                  40                  45

gga gga ggt act ggg tgt cca gtt cac aaa gac aga tct tgt tac att     192
Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile
     50                  55                  60

tgc cac agg cag ctg ctc ttt tgc cgg gta acc ttg gga aag tct ttc     240
Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe
 65                  70                  75                  80

ctg cag ttc agt gca atg aaa atg gca cat tct cct cca ggt cat cac     288
Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His
                 85                  90                  95

tca gtc act ggt agg ccc agt gta aat ggc cta gca tta gct gaa tat     336
Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr
            100                 105                 110

gtt att tac aga gga gaa cag gtaatgtagt tttatttgtt catcttcaaa       387
Val Ile Tyr Arg Gly Glu Gln
        115

aatgctaggg aggcatactt taacttttta ttaatctctt gaattgacaa gacatattgc 447

cttaactgga ttttttaaaa attttatttg gagataattt cagatttgga aagttacaaa 507

aatagtaaag agaattttct tataaccttt acctagattt cctaaatgtt aatattttgt 567

tctctttttt actcttacca ttctctcctt ctttccttgt gtgtgtacct atttttttgt 627

gaactgtttg agagtaagtt gcagggcatg tccctttacc attaactatt tcaattgtaa 687

atttcctaaa aacaagaaga ttttattcaa atttcgccag tcgttccgga tttttcttag 747

ctcttataaa taattgaaat cttgtattta acagcctgtc catagcaaag aagtatataa 807

ctgtgttttg ctctcagtga gagccaaaag tagttctaga gcagtgttgt gaactgggag 867

taggtatcgg aatcaccgca gttactaaaa tcagacatga ttttagtctt atctgatact 927

tatgaactta gtattcatct tagacttgct gattgaaaat ctgaagaact gtactcaggg 987
```

taaagatgtt ttgagaaaat gtccctagat gattctgatc tacaacagta atttagaacc 1047

tcctccctaa gattaggaat acttccggaa agtctgttta tctttcaaga aaatttttgt 1107

accattattt gaatttatct ttctcttcca ggcttatcct gagtatttaa ttacttacca 1167

gattatgagg cctgaaggta tggtcgatgg ataaatagtt attttaagaa actaattcca 1227

ctgaacctaa aatcatcaaa gcagcagtgg cctctacgtt ttactccttt gctgaaaaaa 1287

aatcatcttg cccacaggcc tgtggcaaaa ggataaaaat gtgaacgaag tttaacattc 1347

tgacttgata aagctttaat aatgtacagt gttttctaaa tatttcctgt tttttcagca 1407

ctttaacaga tgccattcca ggttaaactg ggttgtctgt actaaattat aaacagagtt 1467

aacttgaacc ttttatatgt tatgcattga ttctaacaaa ctgtaatgcc ctcaacagaa 1527

ctaattttag taatacaata ctgtgttctt taaaacacag catttacact gaatacaatt 1587

tcatttgtaa aactgtaaat aagagctttt gtactagccc agtatttatt tacattgctt 1647

tgtaatataa tcctgtttta gaagtgcaaa aaaaaaaaaa aaaa 1691

<210> 42
<211> 119
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: Sequence not
specified as protein-coding is vector sequence

<400> 42
Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly
1 5 10 15

Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile
20 25 30

Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile
35 40 45

Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile
50 55 60

Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe
65 70 75 80

Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His
85 90 95

Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr
100 105 110

Val Ile Tyr Arg Gly Glu Gln
115

<210> 43
<211> 1692
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(357)

<220>

<223> Description of Artificial Sequence: Sequence not
specified as protein-coding is vector sequence

<400> 43

```
atg cta ttt cat ggg tct cct ttt gtg aat gca att atc cac aaa ggc    48
Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly
  1               5                  10                  15

ttt gat gaa agg cat gcg tac ata ggt ggt atg ttt gga gct ggc att    96
Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile
             20                  25                  30

tat ttt gct gaa aac tct tcc aaa agc aat caa tat gta tat gga att   144
Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile
         35                  40                  45

gga gga ggt act ggg tgt cca gtt cac aaa gac aga tct tgt tac att   192
Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile
     50                  55                  60

tgc cac agg cag ctg ctc ttt tgc cgg gta acc ttg gga aag tct ttc   240
Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe
 65                  70                  75                  80

ctg cag ttc agt gca atg aaa atg gca cat tct cct cca ggt cat cac   288
Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His
                 85                  90                  95

tca gtc act ggt agg ccc agt gta aat ggc cta gca tta gct gaa tat   336
Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr
            100                 105                 110

gtt att tac aga gga gaa cag gtaatgtagt tttatttgtt catcttcaaa      387
Val Ile Tyr Arg Gly Glu Gln
        115

aatgctaggg aggcatactt taacttttta ttaatctctt gaattgacaa gacatattgc 447

cttaactgga ttttttaaaa attttatttg gagataattt cagatttgga aagttacaaa 507

aatagtaaag agaattttct tataaccttt acctagattt cctaaatgtt aatattttgt 567

tctctttttt actcttacca ttctctcctt ctttccttgt gtgtgtacct atttttttgt 627

gaactgtttg agagtaagtt gcagggcatg tccctttacc attaactatt tcaattgtaa 687

atttcctaaa aacaagaaga ttttattcaa atttcgccag tcgttccgga tttttcttag 747

ctcttataaa taattgaaat cttgtattta acagcctgtc catagcaaag aagtatataa 807

ctgtgttttg ctctcagtga gagccaaaag tagttctaga gcagtgttgt gaactgggag 867

taggtatcgg aatcaccgca gttactaaaa tcagacatga ttttagtctt atctgatact 927

tatgaactta gtattcatct tagacttgct gattgaaaat ctgaagaact gtactcaggg 987

taaagatgtt ttgagaaaat gtccctagat gattctgatc tacaacagta atttagaacc 1047

tcctccctaa gattaggaat acttccggaa agtctgttta tctttcaaga aaatttttgt 1107

accattattt gaatttatct ttctcttcca ggcttatcct gagtatttaa ttacttacca 1167

gattatgagg cctgaaggta tggtcgatgg ataaatagtt attttaagaa actaattcca 1227

ctgaacctaa aatcatcaaa gcagcagtgg cctctacgtt ttactccttt gctgaaaaaa 1287

aatcatcttg cccacaggcc tgtggcaaaa ggataaaaat gtgaacgaag tttaacattc 1347

tgacttgata aagctttaat aatgtacagt gttttctaaa tatttcctgt tttttcagca 1407
```

ctttaacaga tgccattcca ggttaaactg ggttgtctgt actaaattat aaacagagtt 1467

aacttgaacc ttttatatgt tatgcattga ttctaacaaa ctgtaatgcc ctcaacagaa 1527

ctaattttac taatacaata ctgtgttctt taaaacacag catttacact gaatacaatt 1587

tcatttgtaa aactgtaaat aagagctttt gtactagccc agtatttatt tacattgctt 1647

tgtaatataa atctgtttta gaactgcaaa aaaaaaaaaa aaaaa 1692

<210> 44
<211> 119
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: Sequence not
specified as protein-coding is vector sequence

<400> 44
Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly
1 5 10 15

Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile
20 25 30

Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile
35 40 45

Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile
50 55 60

Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe
65 70 75 80

Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His
85 90 95

Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr
100 105 110

Val Ile Tyr Arg Gly Glu Gln
115

<210> 45
<211> 582
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(480)

<220>
<223> Description of Artificial Sequence: Sequence not
specified as protein-coding is vector sequence

<400> 45
gaa aga tac act cac cgg aga aaa gaa gtt tct gaa gaa aac cac aac 48
Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn
1 5 10 15

cat gcc aat gaa cga atg cta ttt cat ggg tct cct ttt gtg aat gca 96
His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala
20 25 30

att atc cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt ggt atg 144
Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met
35 40 45

```
ttt gga gct ggc att tat ttt gct gaa aac tct tcc aaa agc aat caa    192
Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln
     50                  55                  60

tat gta tat gga att gga gga ggt act ggg tgt cca gtt cac aaa gac    240
Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp
 65                  70                  75                  80

aga tct tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg gta acc    288
Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr
                 85                  90                  95

ttg gga aag tct ttc ctg cag ttc agt gca atg aaa atg gca cat tct    336
Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser
            100                 105                 110

cct cca ggt cat cac tca gtc act ggt agg ccc agt gta aat ggc cta    384
Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu
        115                 120                 125

gca tta gct gaa tat gtt att tac aga gga gaa cag gct tat cct gag    432
Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu
    130                 135                 140

tat tta att act tac cag att atg agg cct gaa ggt atg gtc gat gga    480
Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
145                 150                 155                 160

taaatagtta ttttaagaaa ctaattccac tgaacctaaa atcatcaaag cagcagtggc 540

ctctacgttt tactcctttg ctgaaaaaaa aaaaaaaaaa aa                     582
```

<210> 46
<211> 160
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: Sequence not
      specified as protein-coding is vector sequence

<400> 46

```
Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn
  1               5                  10                  15

His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala
             20                  25                  30

Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met
        35                  40                  45

Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln
     50                  55                  60

Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp
 65                  70                  75                  80

Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr
                 85                  90                  95

Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser
            100                 105                 110

Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu
        115                 120                 125

Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu
    130                 135                 140

Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
145                 150                 155                 160
```

<210> 47
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 47
ctccggacaa caaggtctta acc 23

<210> 48
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 48
ccacctatgt acgcatgcc 19

<210> 49
<211> 356
<212> DNA
<213> Homo sapiens

<400> 49
tccggacaac aaggtcttaa cccatattta actttgaaca cctctggtag tggaacaatt 60
cttatagatc tgtctcctga tgataaagag tttcagtctg tggaggaaga gatgcaaagt 120
acagttcgag agcacagaga tggaggtcat gcaggtggaa tcttcaacag atacaatatt 180
ctcaagattc agaaggtttg taacaagaaa ctatgggaaa gatacactca ccggagaaaa 240
gaagtttctg aagaaaacca caaccatgcc aatgaacgaa tgctatttca tgggtctcct 300
tttgtgaatg caattatcca caaaggcttt gatgaaaggc atgcgtacat aggtgg 356

<210> 50
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 50
atttaaccct cactaaaagg g 21

<210> 51
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 51
aaaggctccc atcggcaaat 20

<210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Primer

<400> 52
gttgagggca ttacagtttg 20

<210> 53
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 53
aaaacgtaga ggccactgct 20

<210> 54
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 54
tggtgtagac tgacgccctt 20

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 55
tccggtgagt gtatctttcc 20

<210> 56
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 56
ctcctttgtc ttgggcattc 20

<210> 57
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 57
atctgctctg ccctcttgtt 20

<210> 58
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 58
gggtatcgcg gcaatttaca 20

<210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 59
aacaagaggg cagagcagat 20

<210> 60
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 60
tgccccatct caactaatac 20

<210> 61
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 61
gtaatgccct caacagaact 20

<210> 62
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 62
ggcgtcagtc tacaccactt 20

<210> 63
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 63
taaattgccc gcgataccca 20

<210> 64
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 64
cactcagtca ctggtaggcc 20

<210> 65
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 65
atctgctctg ccctcttgtt 20

<210> 66
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 66
tagttgagat ggggcacaag 20

<210> 67
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 67
aaacgtagag gccactgctg 20

<210> 68
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 68
cgggtaacct tgggaaagtc 20

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 69
gggctttact gctttacaga 20

<210> 70

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 70
gtaagggctg ctgacagtga 20

<210> 71
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 71
ttactccagc agagggcact 20

<210> 72
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 72
ctgacgccct tcaatgtctc 20

<210> 73
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 73
ggtactaagg ccacaattca 20

<210> 74
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 74
gggtatcgcg gcaatttaca 20

<210> 75
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 75
gttgagggca ttacagtttg 20

<210> 76
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 76
taacaagagg gcagagcaga 20

<210> 77
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 77
agttctgttg agggcattac 20

<210> 78
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 78
ggcctaccag tgactgagtg 20

<210> 79
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 79
gggctagagg acctgaagag 20

<210> 80
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 80
agtgccctct gctggagtaa 20

<210> 81
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 81
ggcgtcagtc tacaccactt 20

<210> 82
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 82
tgaattgtgg ccttagtacc 20

<210> 83
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 83
atgcccaaga caaaggagga 20

<210> 84
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 84
gtaatgccct caacagaact 20

<210> 85
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 85
atctgctctg ccctcttgtt 20

<210> 86
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 86
cgggtaacct tgggaaagtc 20

<210> 87
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 87

ccggacaaca aggtcttaac 20

<210> 88
<211> 3353
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2352)

<400> 88
tgt gaa ctg ttg cta aga aaa gga gca aac atc aat gaa aag act aaa 48
Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys
1 5 10 15

gaa ttc ttg act cct ctg cac gtg gca tct gag aaa gct cat aat gat 96
Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn Asp
20 25 30

gtt gtt gaa gta gtg gtg aaa cat gaa gca aag gtt aat gct ctg gat 144
Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu Asp
35 40 45

aat ctt ggt cag act tct cta cac aga gct gca tat tgt ggt cat cta 192
Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu
50 55 60

caa acc tgc cgc cta ctc ctg agc tat ggg tgt gat cct aac att ata 240
Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile
65 70 75 80

tcc ctt cag ggc ttt act gct tta cag atg gga aat gaa aat gta cag 288
Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln
85 90 95

caa ctc ctc caa gag ggt atc tca tta ggt aat tca gag gca gac aga 336
Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg
100 105 110

caa ttg ctg gaa gct gca aag gct gga gat gtc gaa act gta aaa aaa 384
Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys
115 120 125

ctg tgt act gtt cag agt gtc aac tgc aga gac att gaa ggg cgt cag 432
Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln
130 135 140

tct aca cca ctt cat ttt gca gct ggg tat aac aga gtg tcc gtg gtg 480
Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val
145 150 155 160

gaa tat ctg cta cag cat gga gct gat gtg cat gct aaa gat aaa gga 528
Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly
165 170 175

ggc ctt gta cct ttg cac aat gca tgt tct tat gga cat tat gaa gtt 576
Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val
180 185 190

gca gaa ctt ctt gtt aaa cat gga gca gta gtt aat gta gct gat tta 624
Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp Leu
195 200 205

tgg aaa ttt aca cct tta cat gaa gca gca gca aaa gga aaa tat gaa 672
Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu
210 215 220

```
att tgc aaa ctt ctg ctc cag cat ggt gca gac cct aca aaa aaa aac      720
Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn
225                 230                 235                 240

agg gat gga aat act cct ttg gat ctt gtt aaa gat gga gat aca gat      768
Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp
                245                 250                 255

att caa gat ctg ctt agg gga gat gca gct ttg cta gat gct gcc aag      816
Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys
            260                 265                 270

aag ggt tgt tta gcc aga gtg aag aag ttg tct tct cct gat aat gta      864
Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn Val
        275                 280                 285

aat tgc cgc gat acc caa ggc aga cat tca aca cct tta cat tta gca      912
Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu Ala
    290                 295                 300

gct ggt tat aat aat tta gaa gtt gca gag tat ttg tta caa cac gga      960
Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His Gly
305                 310                 315                 320

gct gat gtg aat gcc caa gac aaa gga gga ctt att cct tta cat aat      1008
Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn
                325                 330                 335

gca gca tct tac ggg cat gta gat gta gca gct cta cta ata aag tat      1056
Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys Tyr
            340                 345                 350

aat gca tgt gtc aat gcc acg gac aaa tgg gct ttc aca cct ttg cac      1104
Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His
        355                 360                 365

gaa gca gcc caa aag gga cga aca cag ctt tgt gct ttg ttg cta gcc      1152
Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala
    370                 375                 380

cat gga gct gac ccg act ctt aaa aat cag gaa gga caa aca cct tta      1200
His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro Leu
385                 390                 395                 400

gat tta gtt tca gca gat gat gtc agc gct ctt ctg aca gca gcc atg      1248
Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala Met
                405                 410                 415

ccc cca tct gct ctg ccc tct tgt tac aag cct caa gtg ctc aat ggt      1296
Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn Gly
            420                 425                 430

gtg aga agc cca gga gcc act gca gat gct ctc tct tca ggt cca tct      1344
Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro Ser
        435                 440                 445

agc cca tca agc ctt tct gca gcc agc agt ctt gac aac tta tct ggg      1392
Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser Gly
    450                 455                 460

agt ttt tca gaa ctg tct tca gta gtt agt tca agt gga aca gag ggt      1440
Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu Gly
465                 470                 475                 480

gct tcc agt ttg gag aaa aag gag gtt cca gga gta gat ttt agc ata      1488
Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser Ile
                485                 490                 495
```

```
act caa ttc gta agg aat ctt gga ctt gag cac cta atg gat ata ttt     1536
Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile Phe
            500                 505                 510

gag aga gaa cag atc act ttg gat gta tta gtt gag atg ggg cac aag     1584
Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His Lys
        515                 520                 525

gag ctg aag gag att gga atc aat gct tat gga cat agg cac aaa cta     1632
Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys Leu
    530                 535                 540

att aaa gga gtc gag aga ctt atc tcc gga caa caa ggt ctt aac cca     1680
Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn Pro
545                 550                 555                 560

tat tta act ttg aac acc tct ggt agt gga aca att ctt ata gat ctg     1728
Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp Leu
                565                 570                 575

tct cct gat gat aaa gag ttt cag tct gtg gag gaa gag atg caa agt     1776
Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln Ser
            580                 585                 590

aca gtt cga gag cac aga gat gga ggt cat gca ggt gga atc ttc aac     1824
Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe Asn
        595                 600                 605

aga tac aat att ctc aag att cag aag gtt tgt aac aag aaa cta tgg     1872
Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu Trp
    610                 615                 620

gaa aga tac act cac cgg aga aaa gaa gtt tct gaa gaa aac cac aac     1920
Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn
625                 630                 635                 640

cat gcc aat gaa cga atg cta ttt cat ggg tct cct ttt gtg aat gca     1968
His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala
                645                 650                 655

att atc cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt ggt atg     2016
Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met
            660                 665                 670

ttt gga gct ggc att tat ttt gct gaa aac tct tcc aaa agc aat caa     2064
Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln
        675                 680                 685

tat gta tat gga att gga gga ggt act ggg tgt cca gtt cac aaa gac     2112
Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp
    690                 695                 700

aga tct tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg gta acc     2160
Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr
705                 710                 715                 720

ttg gga aag tct ttc ctg cag ttc agt gca atg aaa atg gca cat tct     2208
Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser
                725                 730                 735

cct cca ggt cat cac tca gtc act ggt agg ccc agt gta aat ggc cta     2256
Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu
            740                 745                 750

gca tta gct gaa tat gtt att tac aga gga gaa cag gct tat cct gag     2304
Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu
        755                 760                 765
```

```
tat tta att act tac cag att atg agg cct gaa ggt atg gtc gat gga    2352
Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
    770                 775                 780

taaatagtta ttttaagaaa ctaattccac tgaacctaaa atcatcaaag cagcagtggc 2412

ctctacgttt tactcctttg ctgaaaaaaa atcatcttgc ccacaggcct gtggcaaaag 2472

gataaaaatg tgaacgaagt ttaacattct gacttgataa agctttaata atgtacagtg 2532

ttttctaaat atttcctgtt ttttcagcac tttaacagat gccattccag gttaaactgg 2592

gttgtctgta ctaaattata aacagagtta acttgaacct tttatatgtt atgcattgat 2652

tctaacaaac tgtaatgccc tcaacagaac taattttact aatacaatac tgtgttcttt 2712

aaaacacagc atttacactg aatacaattt catttgtaaa actgtaaata agagcttttg 2772

tactagccca gtatttattt acattgcttt gtaatataaa tctgttttag aactgcagcg 2832

gtttacaaaa ttttttcata tgtattgttc atctatactt catcttacat cgtcatgatt 2892

gagtgatctt tacatttgat tccagaggct atgttcagtt gttagttggg aaagattgag 2952

ttatcagatt taatttgccg atgggagcct ttatctgtca ttagaaatct ttctcattta 3012

agaacttatg aatatgctga agatttaatt tgtgatacct ttgtatgtat gagacacatt 3072

ccaaagagct ctaactatga taggtcctga ttactaaaga agcttcttta ctggcctcaa 3132

tttctagctt tcatgttgga aaattttctg cagtccttct gtgaaaatta gagcaaagtg 3192

ctcctgtttt ttagagaaac taaatcttgc tgttgaacaa ttattgtgtt cttttcatgg 3252

aacataagta ggatgttaca tttccagggt gggaagggta atcctaaatc atttcccaat 3312

ctattctaat taccttaaat ctaaagggga aaaaaaaaat c                      3353

<210> 89
<211> 784
<212> PRT
<213> Homo sapiens

<400> 89
Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys
1               5                   10                  15

Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn Asp
            20                  25                  30

Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu Asp
        35                  40                  45

Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu
    50                  55                  60

Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile
65                  70                  75                  80

Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln
                85                  90                  95

Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg
            100                 105                 110

Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys
        115                 120                 125
```

```
Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln
    130                 135                 140

Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val
145                 150                 155                 160

Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly
                165                 170                 175

Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val
            180                 185                 190

Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp Leu
        195                 200                 205

Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu
    210                 215                 220

Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn
225                 230                 235                 240

Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp
                245                 250                 255

Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys
            260                 265                 270

Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn Val
        275                 280                 285

Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu Ala
    290                 295                 300

Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His Gly
305                 310                 315                 320

Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn
                325                 330                 335

Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys Tyr
            340                 345                 350

Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His
        355                 360                 365

Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala
    370                 375                 380

His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro Leu
385                 390                 395                 400

Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala Met
                405                 410                 415

Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn Gly
            420                 425                 430

Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro Ser
        435                 440                 445

Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser Gly
    450                 455                 460

Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu Gly
465                 470                 475                 480

Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser Ile
                485                 490                 495
```

```
Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile Phe
            500                 505                 510

Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His Lys
        515                 520                 525

Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys Leu
    530                 535                 540

Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn Pro
545                 550                 555                 560

Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp Leu
                565                 570                 575

Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln Ser
            580                 585                 590

Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe Asn
        595                 600                 605

Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu Trp
    610                 615                 620

Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn
625                 630                 635                 640

His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala
                645                 650                 655

Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met
            660                 665                 670

Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln
        675                 680                 685

Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp
    690                 695                 700

Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr
705                 710                 715                 720

Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser
                725                 730                 735

Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu
            740                 745                 750

Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu
        755                 760                 765

Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
    770                 775                 780


<210> 90
<211> 3799
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(2270)

<400> 90
aa gct cat aat gat gtt gtt gaa gta gtg gtg aaa cat gaa gca aag      47
   Ala His Asn Asp Val Val Glu Val Val Val Lys His Glu Ala Lys
     1               5                  10                  15
```

```
gtt aat gct ctg gat aat ctt ggt cag act tct cta cac aga gct gca     95
Val Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala
                 20                  25                  30

tat tgt ggt cat cta caa acc tgc cgc cta ctc ctg agc tat ggg tgt    143
Tyr Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys
             35                  40                  45

gat cct aac att ata tcc ctt cag ggc ttt act gct tta cag atg gga    191
Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly
         50                  55                  60

aat gaa aat gta cag caa ctc ctc caa gag ggt atc tca tta ggt aat    239
Asn Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn
     65                  70                  75

tca gag gca gac aga caa ttg ctg gaa gct gca aag gct gga gat gtc    287
Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val
 80                  85                  90                  95

gaa act gta aaa aaa ctg tgt act gtt cag agt gtc aac tgc aga gac    335
Glu Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp
                100                 105                 110

att gaa ggg cgt cag tct aca cca ctt cat ttt gca gct ggg tat aac    383
Ile Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn
            115                 120                 125

aga gtg tcc gtg gtg gaa tat ctg cta cag cat gga gct gat gtg cat    431
Arg Val Ser Val Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His
        130                 135                 140

gct aaa gat aaa gga ggc ctt gta cct ttg cac aat gca tgt tct tat    479
Ala Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr
    145                 150                 155

gga cat tat gaa gtt gca gaa ctt ctt gtt aaa cat gga gca gta gtt    527
Gly His Tyr Glu Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val
160                 165                 170                 175

aat gta gct gat tta tgg aaa ttt aca cct tta cat gaa gca gca gca    575
Asn Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala
                180                 185                 190

aaa gga aaa tat gaa att tgc aaa ctt ctg ctc cag cat ggt gca gac    623
Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp
            195                 200                 205

cct aca aaa aaa aac agg gat gga aat act cct ttg gat ctt gtt aaa    671
Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys
        210                 215                 220

gat gga gat aca gat att caa gat ctg ctt agg gga gat gca gct ttg    719
Asp Gly Asp Thr Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu
    225                 230                 235

cta gat gct gcc aag aag ggt tgt tta gcc aga gtg aag aag ttg tct    767
Leu Asp Ala Ala Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser
240                 245                 250                 255

tct cct gat aat gta aat tgc cgc gat acc caa ggc aga cat tca aca    815
Ser Pro Asp Asn Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr
                260                 265                 270

cct tta cat tta gca gct ggt tat aat aat tta gaa gtt gca gag tat    863
Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr
            275                 280                 285
```

ttg tta caa cac gga gct gat gtg aat gcc caa gac aaa gga gga ctt 911
Leu Leu Gln His Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu
290 295 300

att cct tta cat aat gca gca tct tac ggg cat gta gat gta gca gct 959
Ile Pro Leu His Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala
305 310 315

cta cta ata aag tat aat gca tgt gtc aat gcc acg gac aaa tgg gct 1007
Leu Leu Ile Lys Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala
320 325 330 335

ttc aca cct ttg cac gaa gca gcc caa aag gga cga aca cag ctt tgt 1055
Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys
340 345 350

gct ttg ttg cta gcc cat gga gct gac ccg act ctt aaa aat cag gaa 1103
Ala Leu Leu Leu Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu
355 360 365

gga caa aca cct tta gat tta gtt tca gca gat gat gtc agc gct ctt 1151
Gly Gln Thr Pro Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu
370 375 380

ctg aca gca gcc atg ccc cca tct gct ctg ccc tct tgt tac aag cct 1199
Leu Thr Ala Ala Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro
385 390 395

caa gtg ctc aat ggt gtg aga agc cca gga gcc act gca gat gct ctc 1247
Gln Val Leu Asn Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu
400 405 410 415

tct tca ggt cca tct agc cca tca agc ctt tct gca gcc agc agt ctt 1295
Ser Ser Gly Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu
420 425 430

gac aac tta tct ggg agt ttt tca gaa ctg tct tca gta gtt agt tca 1343
Asp Asn Leu Ser Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser
435 440 445

agt gga aca gag ggt gct tcc agt ttg gag aaa aag gag gtt cca gga 1391
Ser Gly Thr Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly
450 455 460

gta gat ttt agc ata act caa ttc gta agg aat ctt gga ctt gag cac 1439
Val Asp Phe Ser Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His
465 470 475

cta atg gat ata ttt gag aga gaa cag atc act ttg gat gta tta gtt 1487
Leu Met Asp Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val
480 485 490 495

gag atg ggg cac aag gag ctg aag gag att gga atc aat gct tat gga 1535
Glu Met Gly His Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly
500 505 510

cat agg cac aaa cta att aaa gga gtc gag aga ctt atc tcc gga caa 1583
His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln
515 520 525

caa ggt ctt aac cca tat tta act ttg aac acc tct ggt agt gga aca 1631
Gln Gly Leu Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr
530 535 540

att ctt ata gat ctg tct cct gat gat aaa gag ttt cag tct gtg gag 1679
Ile Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu
545 550 555

```
gaa gag atg caa agt aca gtt cga gag cac aga gat gga ggt cat gca     1727
Glu Glu Met Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala
560                 565                 570                 575

ggt gga atc ttc aac aga tac aat att ctc aag att cag aag gtt tgt     1775
Gly Gly Ile Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys
                580                 585                 590

aac aag aaa cta tgg gaa aga tac act cac cgg aga aaa gaa gtt tct     1823
Asn Lys Lys Leu Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser
            595                 600                 605

gaa gaa aac cac aac cat gcc aat gaa cga atg cta ttt cat ggg tct     1871
Glu Glu Asn His Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser
        610                 615                 620

cct ttt gtg aat gca att atc cac aaa ggc ttt gat gaa agg cat gcg     1919
Pro Phe Val Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala
    625                 630                 635

tac ata ggt ggt atg ttt gga gct ggc att tat ttt gct gaa aac tct     1967
Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser
640                 645                 650                 655

tcc aaa agc aat caa tat gta tat gga att gga gga ggt act ggg tgt     2015
Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys
                660                 665                 670

cca gtt cac aaa gac aga tct tgt tac att tgc cac agg cag ctg ctc     2063
Pro Val His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu
            675                 680                 685

ttt tgc cgg gta acc ttg gga aag tct ttc ctg cag ttc agt gca atg     2111
Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met
        690                 695                 700

aaa atg gca cat tct cct cca ggt cat cac tca gtc act ggt agg ccc     2159
Lys Met Ala His Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro
    705                 710                 715

agt gta aat ggc cta gca tta gct gaa tat gtt att tac aga gga gaa     2207
Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu
720                 725                 730                 735

cag gct tat cct gag tat tta att act tac cag att atg agg cct gaa     2255
Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu
                740                 745                 750

ggt atg gtc gat gga taaatagtta ttttaagaaa ctaattccac tgaacctaaa     2310
Gly Met Val Asp Gly
            755

atcatcaaag cagcagtggc ctctacgttt tactcctttg ctgaaaaaaa atcatcttgc 2370

ccacaggcct gtggcaaaag gataaaaatg tgaacgaagt ttaacattct gacttgataa 2430

agctttaata atgtacagtg ttttctaaat atttcctgtt ttttcagcac tttaacagat 2490

gccattccag gttaaactgg gttgtctgta ctaaattata aacagagtta acttgaacct 2550

tttatatgtt atgcattgat tctaacaaac tgtaatgccc tcaacagaac taattttact 2610

aatacaatac tgtgttcttt aaaacacagc atttacactg aatacaattt catttgtaaa 2670

actgtaaata agagcttttg tactagccca gtatttattt acattgcttt gtaatataaa 2730

tctgttttag aactgcagcg gtttacaaaa ttttttcata tgtattgttc atctatactt 2790

catcttacat cgtcatgatt gagtgatctt tacatttgat tccagaggct atgttcagtt 2850
```

gttagttggg aaagattgag ttatcagatt taatttgcca ttaaacctta tggggttttc 2910

tgttgcagac tgttgattga ccttactaaa tcccgaaatc taaaaaatga attgtggcct 2970

tagtaccaca ccatctttaa agtctagtgt ttagtcccct tttccttcaa aactttccaa 3030

caaatctagc gctttactga actcagaaca ttgttctctt tgagaatgtg aagattttaa 3090

atagccaaag aattttcatg tataagagct agctaaatat agtatatcct gctctttcga 3150

agaagataca aaactgttgc ctgtactaat gggtatagta gagcagttga agaactaaca 3210

catacatgga cttttcggtc tgaatttgtg ttggcatcca tggtacttac tgttcagtag 3270

gatgttattg caaggagcag agtgccctct gctggagtaa tcgcaattat tcttgcagca 3330

gattaatttg acttgggtca tgaattcaac aaccagttac ttgcctttca tcatacaatt 3390

tcttcggtag ttgagaattt ggtctacatt tatcaaatga ggaaagagtg tcacaaactc 3450

taaaaagctg aaggagaccc cacacatctt ctcactgtca gcagccctta cttctgcaaa 3510

atgttgaagg ataatgtttc tctgtttgca aagaagatgc ctctggctag aatgtttgtg 3570

cagttataag caagggactg cttgtttttg taagttatct caactttatt cttgtgaaat 3630

tgcaaaggaa gatcaataaa aagacttcat ttgaatgtaa atggtgtgaa atactgatgt 3690

gttttgtaca tgtacataat atatttactt cctgctttca cattagtaat ctgagatggt 3750

tctaccattt tataattaga aggagatgta ggggtgggag tggggaggg 3799

<210> 91
<211> 756
<212> PRT
<213> Homo sapiens

<400> 91
Ala His Asn Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val
1 5 10 15

Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr
20 25 30

Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp
35 40 45

Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn
50 55 60

Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser
65 70 75 80

Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu
85 90 95

Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile
100 105 110

Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg
115 120 125

Val Ser Val Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala
130 135 140

Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly
145 150 155 160

His Tyr Glu Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn
165 170 175

Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys
180 185 190

Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro
195 200 205

Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp
210 215 220

Gly Asp Thr Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu
225 230 235 240

Asp Ala Ala Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser
245 250 255

Pro Asp Asn Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro
260 265 270

Leu His Leu Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu
275 280 285

Leu Gln His Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile
290 295 300

Pro Leu His Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu
305 310 315 320

Leu Ile Lys Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe
325 330 335

Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala
340 345 350

Leu Leu Leu Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly
355 360 365

Gln Thr Pro Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu
370 375 380

Thr Ala Ala Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln
385 390 395 400

Val Leu Asn Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser
405 410 415

Ser Gly Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp
420 425 430

Asn Leu Ser Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser
435 440 445

Gly Thr Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val
450 455 460

Asp Phe Ser Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu
465 470 475 480

Met Asp Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu
485 490 495

Met Gly His Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His
500 505 510

Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln
515 520 525

```
Gly Leu Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile
    530                 535                 540

Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu
545                 550                 555                 560

Glu Met Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly
                565                 570                 575

Gly Ile Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn
            580                 585                 590

Lys Lys Leu Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu
        595                 600                 605

Glu Asn His Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro
    610                 615                 620

Phe Val Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr
625                 630                 635                 640

Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser
                645                 650                 655

Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro
            660                 665                 670

Val His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe
        675                 680                 685

Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys
    690                 695                 700

Met Ala His Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser
705                 710                 715                 720

Val Asn Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln
                725                 730                 735

Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly
            740                 745                 750

Met Val Asp Gly
        755
```

<210> 92
<211> 2971
<212> DNA
<213> Homo sapiens

<400> 92

```
tgtgaactgt tgctaagaaa aggagcaaac atcaatgaaa agactaaaga attcttgact 60
cctctgcacg tggcatctga gaaagctcat aatgatgttg ttgaagtagt ggtgaaacat 120
gaagcaaagg ttaatgctct ggataatctt ggtcagactt ctctacacag agctgcatat 180
tgtggtcatc tacaaacctg ccgcctactc ctgagctatg ggtgtgatcc taacattata 240
tcccttcagg gctttactgc tttacagatg ggaaatgaaa atgtacagca actcctccaa 300
gagggtatct cattaggtaa ttcagaggca gacagacaat tgctggaagc tgcaaaggct 360
ggagatgtcg aaactgtaaa aaaactgtgt actgttcaga gtgtcaactg cagagacatt 420
gaagggcgtc agtctacacc acttcatttt gcagctgggt ataacagagt gtccgtggtg 480
gaatatctgc tacagcatgg agctgatgtg catgctaaag ataaaggagg ccttgtacct 540
ttgcacaatg catgttctta tggacattat gaagttgcag aacttcttgt taaacatgga 600
gcagtagtta atgtagctga tttatggaaa tttacacctt tacatgaagc agcagcaaaa 660
ggaaaatatg aaatttgcaa acttctgctc cagcatggtg cagaccctac aaaaaaaaac 720
agggatggaa atactccttt ggatcttgtt aaagatggag atacagatat tcaagatctg 780
cttaggggag atgcagcttt gctagatgct gccaagaagg gttgtttagc cagagtgaag 840
aagttgtctt ctcctgataa tgtaaattgc cgcgataccc aaggcagaca ttcaacacct 900
ttacatttag cagctggtta taataattta gaagttgcag agtatttgtt acaacacgga 960
gctgatgtga atgcccaaga caaaggagga cttattcctt tacataatgc agcatcttac 1020
```

gggcatgtag atgtagcagc tctactaata aagtataatg catgtgtcaa tgccacggac 1080
aaatgggctt tcacaccttt gcacgaagca gcccaaaagg gacgaacaca gctttgtgct 1140
ttgttgctag cccatggagc tgacccgact cttaaaaatc aggaaggaca aacaccttta 1200
gatttagttt cagcagatga tgtcagcgct cttctgacag cagccatgcc cccatctgct 1260
ctgccctctt gttacaagcc tcaagtgctc aatggtgtga gaagcccagg agccactgca 1320
gatgctctct cttcaggtcc atctagccca tcaagccttt ctgcagccag cagtcttgac 1380
aacttatctg ggagtttttc agaactgtct tcagtagtta gttcaagtgg aacagagggt 1440
gcttccagtt tggagaaaaa ggaggttcca ggagtagatt ttagcataac tcaattcgta 1500
aggaatcttg gacttgagca cctaatggat atatttgaga gagaacagat cactttggat 1560
gtattagttg agatggggca caaggagctg aaggagattg gaatcaatgc ttatggacat 1620
aggcacaaac taattaaagg agtcgagaga cttatctccg gacaacaagg tcttaaccca 1680
tatttaactt tgaacacctc tggtagtgga acaattctta tagatctgtc tcctgatgat 1740
aaagagtttc agtctgtgga ggaagagatg caaagtacag ttcgagagca cagagatgga 1800
ggtcatgcag gtggaatctt caacagatac aatattctca agattcagaa ggtttgtaac 1860
aagaaactat gggaaagata cactcaccgg agaaaagaag tttctgaaga aaaccacaac 1920
catgccaatg aacgaatgct atttcatggg tctccttttg tgaatgcaat tatccacaaa 1980
ggctttgatg aaaggcatgc gtacataggt ggtatgtttg gagctggcat ttattttgct 2040
gaaaactctt ccaaaagcaa tcaatatgta tatggaattg gaggaggtac tgggtgtcca 2100
gttcacaaag acagatcttg ttacatttgc cacaggcagc tgctcttttg ccgggtaacc 2160
ttgggaaagt ctttcctgca gttcagtgca atgaaaatgg cacattctcc tccaggtcat 2220
cactcagtca ctggtaggcc cagtgtaaat ggcctagcat tagctgaata tgttatttac 2280
agaggagaac aggcttatcc tgagtattta attacttacc agattatgag gcctgaaggt 2340
atggtcgatg gataaatagt tattttaaga aactaattcc actgaaccta aaatcatcaa 2400
agcagcagtg gcctctacgt tttactcctt tgctgaaaaa aaatcatctt gcccacaggc 2460
ctgtggcaaa aggataaaaa tgtgaacgaa gtttaacatt ctgacttgat aaagctttaa 2520
taatgtacag tgttttctaa atatttcctg ttttttcagc actttaacag atgccattcc 2580
aggttaaact gggttgtctg tactaaatta taaacagagt taacttgaac cttttatatg 2640
ttatgcattg attctaacaa actgtaatgc cctcaacaga actaatttta ctaatacaat 2700
actgtgttct ttaaaacaca gcatttacac tgaatacaat ttcatttgta aaactgtaaa 2760
taagagcttt tgtactagcc cagtatttat ttacattgct ttgtaatata aatctgtttt 2820
agaactgcag cggtttacaa aattttttca tatgtattgt tcatctatac ttcatcttac 2880
atcgtcatga ttgagtgatc tttacatttg attccagagg ctatgttcag ttgttagttg 2940
ggaaagattg agttatcaga tttaatttgc c 2971

<210> 93
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 93
gggcggaaag acgtagttga 20

<210> 94
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 94
gcggctgttc accttctcag 20

<210> 95
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 95
acgcaagtga tggcagaaag 20

<210> 96
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 96
tcacttgcgt ggcagttgac 20

<210> 97
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 97
gcggcaggtt tgtagatgac 20

<210> 98
<211> 1568
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (2)..(1567)

<400> 98
```
g gcc agg atc atg tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc 49
  Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
    1               5                  10                  15

gcg agc gcc gcg gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc   97
Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
                20                  25                  30

gag gcg tgc cgc aac ggg gac gtg gaa cga gtc aag agg ctg gtg acg   145
Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
            35                  40                  45

cct gag aag gtg aac agc cgc gac acg gcg ggc agg aaa tcc acc ccg   193
Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
    50                  55                  60

ctg cac ttc gcc gca ggt ttt ggg cgg aaa gac gta gtt gaa tat ttg   241
Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
65                  70                  75                  80

ctt cag aat ggt gca aat gtc caa gca cgt gat gat ggg ggc ctt att   289
Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
                85                  90                  95

cct ctt cat aat gca tgc tct ttt ggt cat gct gaa gta gtc aat ctc   337
Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
            100                 105                 110

ctt ttg cga cat ggt gca gac ccc aat gct cga gat aat tgg aat tat   385
Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
        115                 120                 125

act cct ctc cat gaa gct gca att aaa gga aag att gat gtt tgc att   433
Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
    130                 135                 140
```

```
gtg ctg tta cag cat gga gct gag cca acc atc cga aat aca gat gga    481
Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145                 150                 155                 160

agg aca gca ttg gat tta gca gat cca tct gcc aaa gca gtg ctt act    529
Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
                165                 170                 175

ggt gaa tat aag aaa gat gaa ctc tta gaa agt gcc agg agt ggc aat    577
Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
            180                 185                 190

gaa gaa aaa atg atg gct cta ctc aca cca tta aat gtc aac tgc cac    625
Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
        195                 200                 205

gca agt gat ggc aga aag tca act cca tta cat ttg gca gca gga tat    673
Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
    210                 215                 220

aac aga gta aag att gta cag ctg tta ctg caa cat gga gct gat gtc    721
Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225                 230                 235                 240

cat gct aaa gat aaa ggt gat ctg gta cca tta cac aat gcc tgt tct    769
His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
                245                 250                 255

tat ggt cat tat gaa gta act gaa ctt ttg gtc aag cat ggt gcc tgt    817
Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
            260                 265                 270

gta aat gca atg gac ttg tgg caa ttc act cct ctt cat gag gca gct    865
Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
        275                 280                 285

tct aag aac agg gtt gaa gta tgt tct ctt ctc tta agt tat ggt gca    913
Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
    290                 295                 300

gac cca aca ctg ctc aat tgt cac aat aaa agt gct ata gac ttg gct    961
Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305                 310                 315                 320

ccc aca cca cag tta aaa gaa aga tta gca tat gaa ttt aaa ggc cac    1009
Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
                325                 330                 335

tcg ttg ctg caa gct gca cga gaa gct gat gtt act cga atc aaa aaa    1057
Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
            340                 345                 350

cat ctc tct ctg gaa atg gtg aat ttc aag cat cct caa aca cat gaa    1105
His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
        355                 360                 365

aca gca ttg cat tgt gct gct gca tct cca tat ccc aaa aga aag caa    1153
Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
    370                 375                 380

ata tgt gaa ctg ttg cta aga aaa gga gca aac atc aat gaa aag act    1201
Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385                 390                 395                 400

aaa gaa ttc ttg act cct ctg cac gtg gca tct gag aaa gct cat aat    1249
Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
                405                 410                 415
```

```
gat gtt gtt gaa gta gtg gtg aaa cat gaa gca aag gtt aat gct ctg    1297
Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
            420                 425                 430

gat aat ctt ggt cag act tct cta cac aga gct gca tat tgt ggt cat    1345
Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
        435                 440                 445

cta caa acc tgc cgc cta ctc ctg agc tat ggg tgt gat cct aac att    1393
Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
    450                 455                 460

ata tcc ctt cag ggc ttt act gct tta cag atg gga aat gaa aat gta    1441
Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465                 470                 475                 480

cag caa ctc ctc caa gag ggt atc tca tta ggt aat tca gag gca gac    1489
Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
                485                 490                 495

aga caa ttg ctg gaa gct gca aag gct gga gat gtc gaa act gta aaa    1537
Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
            500                 505                 510

aaa ctg tgt act gtt cag agt gtc aac tgc a                          1568
Lys Leu Cys Thr Val Gln Ser Val Asn Cys
        515                 520


<210> 99
<211> 522
<212> PRT
<213> Homo sapiens

<400> 99
Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
1               5                   10                  15

Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
            20                  25                  30

Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
        35                  40                  45

Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
    50                  55                  60

Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
65                  70                  75                  80

Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
                85                  90                  95

Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
            100                 105                 110

Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
        115                 120                 125

Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
    130                 135                 140

Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145                 150                 155                 160

Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
                165                 170                 175

Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
            180                 185                 190
```

Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
195 200 205

Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
210 215 220

Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225 230 235 240

His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
245 250 255

Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
260 265 270

Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
275 280 285

Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
290 295 300

Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305 310 315 320

Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
325 330 335

Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
340 345 350

His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
355 360 365

Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
370 375 380

Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385 390 395 400

Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
405 410 415

Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
420 425 430

Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
435 440 445

Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
450 455 460

Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465 470 475 480

Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
485 490 495

Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
500 505 510

Lys Leu Cys Thr Val Gln Ser Val Asn Cys
515 520

<210> 100
<211> 4127
<212> DNA
<213> Homo sapiens

```
<220>
<221> CDS
<222> (2)..(3508)

<220>
<221> 3'UTR
<222> (3509)..(4127)

<400> 100
g gcc agg atc atg tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc 49
  Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
    1               5                  10                  15

gcg agc gcc gcg gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc    97
Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
            20                  25                  30

gag gcg tgc cgc aac ggg gac gtg gaa cga gtc aag agg ctg gtg acg    145
Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
        35                  40                  45

cct gag aag gtg aac agc cgc gac acg gcg ggc agg aaa tcc acc ccg    193
Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
    50                  55                  60

ctg cac ttc gcc gca ggt ttt ggg cgg aaa gac gta gtt gaa tat ttg    241
Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
65                  70                  75                  80

ctt cag aat ggt gca aat gtc caa gca cgt gat gat ggg ggc ctt att    289
Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
                85                  90                  95

cct ctt cat aat gca tgc tct ttt ggt cat gct gaa gta gtc aat ctc    337
Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
            100                 105                 110

ctt ttg cga cat ggt gca gac ccc aat gct cga gat aat tgg aat tat    385
Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
        115                 120                 125

act cct ctc cat gaa gct gca att aaa gga aag att gat gtt tgc att    433
Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
    130                 135                 140

gtg ctg tta cag cat gga gct gag cca acc atc cga aat aca gat gga    481
Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145                 150                 155                 160

agg aca gca ttg gat tta gca gat cca tct gcc aaa gca gtg ctt act    529
Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
                165                 170                 175

ggt gaa tat aag aaa gat gaa ctc tta gaa agt gcc agg agt ggc aat    577
Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
            180                 185                 190

gaa gaa aaa atg atg gct cta ctc aca cca tta aat gtc aac tgc cac    625
Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
        195                 200                 205

gca agt gat ggc aga aag tca act cca tta cat ttg gca gca gga tat    673
Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
    210                 215                 220

aac aga gta aag att gta cag ctg tta ctg caa cat gga gct gat gtc    721
Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225                 230                 235                 240
```

```
cat gct aaa gat aaa ggt gat ctg gta cca tta cac aat gcc tgt tct     769
His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
            245                 250                 255

tat ggt cat tat gaa gta act gaa ctt ttg gtc aag cat ggt gcc tgt     817
Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
        260                 265                 270

gta aat gca atg gac ttg tgg caa ttc act cct ctt cat gag gca gct     865
Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
    275                 280                 285

tct aag aac agg gtt gaa gta tgt tct ctt ctc tta agt tat ggt gca     913
Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
290                 295                 300

gac cca aca ctg ctc aat tgt cac aat aaa agt gct ata gac ttg gct     961
Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305                 310                 315                 320

ccc aca cca cag tta aaa gaa aga tta gca tat gaa ttt aaa ggc cac     1009
Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
                325                 330                 335

tcg ttg ctg caa gct gca cga gaa gct gat gtt act cga atc aaa aaa     1057
Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
            340                 345                 350

cat ctc tct ctg gaa atg gtg aat ttc aag cat cct caa aca cat gaa     1105
His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
        355                 360                 365

aca gca ttg cat tgt gct gct gca tct cca tat ccc aaa aga aag caa     1153
Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
    370                 375                 380

ata tgt gaa ctg ttg cta aga aaa gga gca aac atc aat gaa aag act     1201
Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385                 390                 395                 400

aaa gaa ttc ttg act cct ctg cac gtg gca tct gag aaa gct cat aat     1249
Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
                405                 410                 415

gat gtt gtt gaa gta gtg gtg aaa cat gaa gca aag gtt aat gct ctg     1297
Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
            420                 425                 430

gat aat ctt ggt cag act tct cta cac aga gct gca tat tgt ggt cat     1345
Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
        435                 440                 445

cta caa acc tgc cgc cta ctc ctg agc tat ggg tgt gat cct aac att     1393
Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
    450                 455                 460

ata tcc ctt cag ggc ttt act gct tta cag atg gga aat gaa aat gta     1441
Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465                 470                 475                 480

cag caa ctc ctc caa gag ggt atc tca tta ggt aat tca gag gca gac     1489
Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
                485                 490                 495

aga caa ttg ctg gaa gct gca aag gct gga gat gtc gaa act gta aaa     1537
Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
            500                 505                 510
```

aaa ctg tgt act gtt cag agt gtc aac tgc aga gac att gaa ggg cgt 1585
Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg
515 520 525

cag tct aca cca ctt cat ttt gca gct ggg tat aac aga gtg tcc gtg 1633
Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val
530 535 540

gtg gaa tat ctg cta cag cat gga gct gat gtg cat gct aaa gat aaa 1681
Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
545 550 555 560

gga ggc ctt gta cct ttg cac aat gca tgt tct tat gga cat tat gaa 1729
Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
565 570 575

gtt gca gaa ctt ctt gtt aaa cat gga gca gta gtt aat gta gct gat 1777
Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp
580 585 590

tta tgg aaa ttt aca cct tta cat gaa gca gca gca aaa gga aaa tat 1825
Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr
595 600 605

gaa att tgc aaa ctt ctg ctc cag cat ggt gca gac cct aca aaa aaa 1873
Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys
610 615 620

aac agg gat gga aat act cct ttg gat ctt gtt aaa gat gga gat aca 1921
Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr
625 630 635 640

gat att caa gat ctg ctt agg gga gat gca gct ttg cta gat gct gcc 1969
Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala
645 650 655

aag aag ggt tgt tta gcc aga gtg aag aag ttg tct tct cct gat aat 2017
Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn
660 665 670

gta aat tgc cgc gat acc caa ggc aga cat tca aca cct tta cat tta 2065
Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu
675 680 685

gca gct ggt tat aat aat tta gaa gtt gca gag tat ttg tta caa cac 2113
Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His
690 695 700

gga gct gat gtg aat gcc caa gac aaa gga gga ctt att cct tta cat 2161
Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His
705 710 715 720

aat gca gca tct tac ggg cat gta gat gta gca gct cta cta ata aag 2209
Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys
725 730 735

tat aat gca tgt gtc aat gcc acg gac aaa tgg gct ttc aca cct ttg 2257
Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu
740 745 750

cac gaa gca gcc caa aag gga cga aca cag ctt tgt gct ttg ttg cta 2305
His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu
755 760 765

gcc cat gga gct gac ccg act ctt aaa aat cag gaa gga caa aca cct 2353
Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro
770 775 780

```
tta gat tta gtt tca gca gat gat gtc agc gct ctt ctg aca gca gcc    2401
Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala
785                 790                 795                 800

atg ccc cca tct gct ctg ccc tct tgt tac aag cct caa gtg ctc aat    2449
Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn
            805                 810                 815

ggt gtg aga agc cca gga gcc act gca gat gct ctc tct tca ggt cca    2497
Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro
        820                 825                 830

tct agc cca tca agc ctt tct gca gcc agc agt ctt gac aac tta tct    2545
Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser
    835                 840                 845

ggg agt ttt tca gaa ctg tct tca gta gtt agt tca agt gga aca gag    2593
Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu
850                 855                 860

ggt gct tcc agt ttg gag aaa aag gag gtt cca gga gta gat ttt agc    2641
Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser
865                 870                 875                 880

ata act caa ttc gta agg aat ctt gga ctt gag cac cta atg gat ata    2689
Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile
            885                 890                 895

ttt gag aga gaa cag atc act ttg gat gta tta gtt gag atg ggg cac    2737
Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His
        900                 905                 910

aag gag ctg aag gag att gga atc aat gct tat gga cat agg cac aaa    2785
Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys
    915                 920                 925

cta att aaa gga gtc gag aga ctt atc tcc gga caa caa ggt ctt aac    2833
Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn
930                 935                 940

cca tat tta act ttg aac acc tct ggt agt gga aca att ctt ata gat    2881
Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp
945                 950                 955                 960

ctg tct cct gat gat aaa gag ttt cag tct gtg gag gaa gag atg caa    2929
Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln
            965                 970                 975

agt aca gtt cga gag cac aga gat gga ggt cat gca ggt gga atc ttc    2977
Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe
        980                 985                 990

aac aga tac aat att ctc aag att cag aag gtt tgt aac aag aaa cta    3025
Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu
    995                 1000                1005

tgg gaa aga tac act cac cgg aga aaa gaa gtt tct gaa gaa aac cac    3073
Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His
    1010                1015                1020

aac cat gcc aat gaa cga atg cta ttt cat ggg tct cct ttt gtg aat    3121
Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn
1025                1030                1035                1040

gca att atc cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt ggt    3169
Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly
                1045                1050                1055
```

```
atg ttt gga gct ggc att tat ttt gct gaa aac tct tcc aaa agc aat    3217
Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn
            1060                1065                1070

caa tat gta tat gga att gga gga ggt act ggg tgt cca gtt cac aaa    3265
Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys
        1075                1080                1085

gac aga tct tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg gta    3313
Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val
    1090                1095                1100

acc ttg gga aag tct ttc ctg cag ttc agt gca atg aaa atg gca cat    3361
Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His
1105                1110                1115                1120

tct cct cca ggt cat cac tca gtc act ggt agg ccc agt gta aat ggc    3409
Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly
                1125                1130                1135

cta gca tta gct gaa tat gtt att tac aga gga gaa cag gct tat cct    3457
Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro
            1140                1145                1150

gag tat tta att act tac cag att atg agg cct gaa ggt atg gtc gat    3505
Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp
        1155                1160                1165

gga taaatagtta ttttaagaaa ctaattccac tgaacctaaa atcatcaaag         3558
Gly

cagcagtggc ctctacgttt tactcctttg ctgaaaaaaa atcatcttgc ccacaggcct 3618

gtggcaaaag gataaaaatg tgaacgaagt ttaacattct gacttgataa agctttaata 3678

atgtacagtg ttttctaaat atttcctgtt ttttcagcac tttaacagat gccattccag 3738

gttaaactgg gttgtctgta ctaaattata aacagagtta acttgaacct tttatatgtt 3798

atgcattgat tctaacaaac tgtaatgccc tcaacagaac taattttact aatacaatac 3858

tgtgttcttt aaaacacagc atttacactg aatacaattt catttgtaaa actgtaaata 3918

agagcttttg tactagccca gtatttattt acattgcttt gtaatataaa tctgttttag 3978

aactgcagcg gtttacaaaa ttttttcata tgtattgttc atctatactt catcttacat 4038

cgtcatgatt gagtgatctt tacatttgat tccagaggct atgttcagtt gttagttggg 4098

aaagattgag ttatcagatt taatttgcc                                    4127


<210> 101
<211> 1169
<212> PRT
<213> Homo sapiens

<400> 101
Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys
  1               5                  10                  15

Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe
                20                  25                  30

Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr
        35                  40                  45

Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro
    50                  55                  60
```

```
Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu
 65                  70                  75                  80

Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile
                 85                  90                  95

Pro Leu His Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu
            100                 105                 110

Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr
        115                 120                 125

Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile
    130                 135                 140

Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly
145                 150                 155                 160

Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr
                165                 170                 175

Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn
            180                 185                 190

Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His
        195                 200                 205

Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr
    210                 215                 220

Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val
225                 230                 235                 240

His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser
                245                 250                 255

Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys
            260                 265                 270

Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala
        275                 280                 285

Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala
    290                 295                 300

Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala
305                 310                 315                 320

Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His
                325                 330                 335

Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys
            340                 345                 350

His Leu Ser Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu
        355                 360                 365

Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln
    370                 375                 380

Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr
385                 390                 395                 400

Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn
                405                 410                 415

Asp Val Val Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu
            420                 425                 430
```

Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His
435 440 445

Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile
450 455 460

Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val
465 470 475 480

Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp
485 490 495

Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys
500 505 510

Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg
515 520 525

Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val
530 535 540

Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys
545 550 555 560

Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu
565 570 575

Val Ala Glu Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp
580 585 590

Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr
595 600 605

Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys
610 615 620

Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr
625 630 635 640

Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala
645 650 655

Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn
660 665 670

Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu
675 680 685

Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His
690 695 700

Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His
705 710 715 720

Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys
725 730 735

Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu
740 745 750

His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu
755 760 765

Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro
770 775 780

Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala
785 790 795 800

```
Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn
                805                 810                 815

Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro
            820                 825                 830

Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser
        835                 840                 845

Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu
    850                 855                 860

Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser
865                 870                 875                 880

Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile
                885                 890                 895

Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His
            900                 905                 910

Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys
        915                 920                 925

Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn
    930                 935                 940

Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp
945                 950                 955                 960

Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln
                965                 970                 975

Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe
            980                 985                 990

Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu
        995                 1000                1005

Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His
    1010                1015                1020

Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn
1025                1030                1035                1040

Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly
                1045                1050                1055

Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn
            1060                1065                1070

Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys
        1075                1080                1085

Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val
    1090                1095                1100

Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His
1105                1110                1115                1120

Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly
                1125                1130                1135

Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro
            1140                1145                1150

Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp
        1155                1160                1165

Gly
```

<210> 102
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 102
gagcattggg gtctgcacca tgtcgcaaaa gg 32

<210> 103
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 103
ccatcctaat acgactcact atagggc 27

<210> 104
<211> 647
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (2)..(646)

<400> 104
g gag ctg gca gga ggg gcc ttg cca gct tcc gcc gcc gcg tcg ttt cag 49
Glu Leu Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln
1 5 10 15

gac ccg gac ggc gga ttc gcg ctg cct ccg ccg ccg cgg ggc agc cgg 97
Asp Pro Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg
20 25 30

ggg gca ggg agc cca gcg agg ggc gcg cgt ggg cgc ggc cat ggg act 145
Gly Ala Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr
35 40 45

gcg ccg gat ccg gtg aca gca ggg agc caa gcg gcc cgg gcc ctg agc 193
Ala Pro Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser
50 55 60

gcg tct tct ccg ggg ggc ctc gcc ctc ctg ctc gcg ggg ccg ggg ctc 241
Ala Ser Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu
65 70 75 80

ctg ctc cgg ttg ctg gcg ctg ttg ctg gct gtg gcg gcg gcc agg atc 289
Leu Leu Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile
85 90 95

atg tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc gcg agc gcc 337
Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala
100 105 110

gcg gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc gag gcg tgc 385
Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
115 120 125

```
cgc aac ggg gac gtg gaa cga gtc aag agg ctg gtg acg cct gag aag     433
Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys
    130                 135                 140

gtg aac agc cgc gac acg gcg ggc agg aaa tcc acc ccg ctg cac ttc     481
Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe
145                 150                 155                 160

gcc gca ggt ttt ggg cgg aaa gac gta gtt gaa tat ttg ctt cag aat     529
Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn
                165                 170                 175

ggt gca aat gtc caa gca cgt gat gat ggg ggc ctt att cct ctt cat     577
Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His
            180                 185                 190

aat gca tgc tct ttt ggt cat gct gaa gta gtc aat ctc ctt ttg cga     625
Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu Leu Leu Arg
        195                 200                 205

cat ggt gca gac ccc aat gct c                                       647
His Gly Ala Asp Pro Asn Ala
    210                 215


&lt;210&gt; 105
&lt;211&gt; 215
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 105
Glu Leu Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln
 1               5                  10                  15

Asp Pro Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg
            20                  25                  30

Gly Ala Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr
        35                  40                  45

Ala Pro Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser
    50                  55                  60

Ala Ser Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu
65                  70                  75                  80

Leu Leu Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile
                85                  90                  95

Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala
            100                 105                 110

Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
        115                 120                 125

Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys
    130                 135                 140

Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe
145                 150                 155                 160

Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn
                165                 170                 175

Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His
            180                 185                 190

Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu Leu Leu Arg
        195                 200                 205
```

```
His Gly Ala Asp Pro Asn Ala
    210                 215


<210> 106
<211> 4406
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (2)..(3787)

<400> 106
g gag ctg gca gga ggg gcc ttg cca gct tcc gcc gcc gcg tcg ttt cag 49
  Glu Leu Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln
    1               5                   10                  15

gac ccg gac ggc gga ttc gcg ctg cct ccg ccg ccg cgg ggc agc cgg    97
Asp Pro Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg
                20                  25                  30

ggg gca ggg agc cca gcg agg ggc gcg cgt ggg cgc ggc cat ggg act    145
Gly Ala Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr
            35                  40                  45

gcg ccg gat ccg gtg aca gca ggg agc caa gcg gcc cgg gcc ctg agc    193
Ala Pro Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser
        50                  55                  60

gcg tct tct ccg ggg ggc ctc gcc ctc ctg ctc gcg ggg ccg ggg ctc    241
Ala Ser Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu
 65                  70                  75                  80

ctg ctc cgg ttg ctg gcg ctg ttg ctg gct gtg gcg gcg gcc agg atc    289
Leu Leu Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile
                85                  90                  95

atg tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc gcg agc gcc    337
Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala
            100                 105                 110

gcg gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc gag gcg tgc    385
Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
        115                 120                 125

cgc aac ggg gac gtg gaa cga gtc aag agg ctg gtg acg cct gag aag    433
Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys
    130                 135                 140

gtg aac agc cgc gac acg gcg ggc agg aaa tcc acc ccg ctg cac ttc    481
Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe
145                 150                 155                 160

gcc gca ggt ttt ggg cgg aaa gac gta gtt gaa tat ttg ctt cag aat    529
Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn
                165                 170                 175

ggt gca aat gtc caa gca cgt gat gat ggg ggc ctt att cct ctt cat    577
Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His
            180                 185                 190

aat gca tgc tct ttt ggt cat gct gaa gta gtc aat ctc ctt ttg cga    625
Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu Leu Leu Arg
        195                 200                 205

cat ggt gca gac ccc aat gct cga gat aat tgg aat tat act cct ctc    673
His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu
    210                 215                 220
```

```
cat gaa gct gca att aaa gga aag att gat gtt tgc att gtg ctg tta     721
His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu
225                 230                 235                 240

cag cat gga gct gag cca acc atc cga aat aca gat gga agg aca gca     769
Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly Arg Thr Ala
                245                 250                 255

ttg gat tta gca gat cca tct gcc aaa gca gtg ctt act ggt gaa tat     817
Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr
            260                 265                 270

aag aaa gat gaa ctc tta gaa agt gcc agg agt ggc aat gaa gaa aaa     865
Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn Glu Glu Lys
        275                 280                 285

atg atg gct cta ctc aca cca tta aat gtc aac tgc cac gca agt gat     913
Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp
    290                 295                 300

ggc aga aag tca act cca tta cat ttg gca gca gga tat aac aga gta     961
Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val
305                 310                 315                 320

aag att gta cag ctg tta ctg caa cat gga gct gat gtc cat gct aaa     1009
Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val His Ala Lys
                325                 330                 335

gat aaa ggt gat ctg gta cca tta cac aat gcc tgt tct tat ggt cat     1057
Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His
            340                 345                 350

tat gaa gta act gaa ctt ttg gtc aag cat ggt gcc tgt gta aat gca     1105
Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys Val Asn Ala
        355                 360                 365

atg gac ttg tgg caa ttc act cct ctt cat gag gca gct tct aag aac     1153
Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn
    370                 375                 380

agg gtt gaa gta tgt tct ctt ctc tta agt tat ggt gca gac cca aca     1201
Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr
385                 390                 395                 400

ctg ctc aat tgt cac aat aaa agt gct ata gac ttg gct ccc aca cca     1249
Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro
                405                 410                 415

cag tta aaa gaa aga tta gca tat gaa ttt aaa ggc cac tcg ttg ctg     1297
Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu
            420                 425                 430

caa gct gca cga gaa gct gat gtt act cga atc aaa aaa cat ctc tct     1345
Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser
        435                 440                 445

ctg gaa atg gtg aat ttc aag cat cct caa aca cat gaa aca gca ttg     1393
Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu Thr Ala Leu
    450                 455                 460

cat tgt gct gct gca tct cca tat ccc aaa aga aag caa ata tgt gaa     1441
His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu
465                 470                 475                 480

ctg ttg cta aga aaa gga gca aac atc aat gaa aag act aaa gaa ttc     1489
Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe
                485                 490                 495
```

```
ttg act cct ctg cac gtg gca tct gag aaa gct cat aat gat gtt gtt    1537
Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn Asp Val Val
            500                 505                 510

gaa gta gtg gtg aaa cat gaa gca aag gtt aat gct ctg gat aat ctt    1585
Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu
        515                 520                 525

ggt cag act tct cta cac aga gct gca tat tgt ggt cat cta caa acc    1633
Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr
    530                 535                 540

tgc cgc cta ctc ctg agc tat ggg tgt gat cct aac att ata tcc ctt    1681
Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu
545                 550                 555                 560

cag ggc ttt act gct tta cag atg gga aat gaa aat gta cag caa ctc    1729
Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu
                565                 570                 575

ctc caa gag ggt atc tca tta ggt aat tca gag gca gac aga caa ttg    1777
Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu
            580                 585                 590

ctg gaa gct gca aag gct gga gat gtc gaa act gta aaa aaa ctg tgt    1825
Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys
        595                 600                 605

act gtt cag agt gtc aac tgc aga gac att gaa ggg cgt cag tct aca    1873
Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr
    610                 615                 620

cca ctt cat ttt gca gct ggg tat aac aga gtg tcc gtg gtg gaa tat    1921
Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr
625                 630                 635                 640

ctg cta cag cat gga gct gat gtg cat gct aaa gat aaa gga ggc ctt    1969
Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu
                645                 650                 655

gta cct ttg cac aat gca tgt tct tat gga cat tat gaa gtt gca gaa    2017
Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu
            660                 665                 670

ctt ctt gtt aaa cat gga gca gta gtt aat gta gct gat tta tgg aaa    2065
Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys
        675                 680                 685

ttt aca cct tta cat gaa gca gca gca aaa gga aaa tat gaa att tgc    2113
Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys
    690                 695                 700

aaa ctt ctg ctc cag cat ggt gca gac cct aca aaa aaa aac agg gat    2161
Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp
705                 710                 715                 720

gga aat act cct ttg gat ctt gtt aaa gat gga gat aca gat att caa    2209
Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp Ile Gln
                725                 730                 735

gat ctg ctt agg gga gat gca gct ttg cta gat gct gcc aag aag ggt    2257
Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly
            740                 745                 750

tgt tta gcc aga gtg aag aag ttg tct tct cct gat aat gta aat tgc    2305
Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn Val Asn Cys
        755                 760                 765
```

```
cgc gat acc caa ggc aga cat tca aca cct tta cat tta gca gct ggt    2353
Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu Ala Ala Gly
    770                 775                 780

tat aat aat tta gaa gtt gca gag tat ttg tta caa cac gga gct gat    2401
Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His Gly Ala Asp
785                 790                 795                 800

gtg aat gcc caa gac aaa gga gga ctt att cct tta cat aat gca gca    2449
Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala
                805                 810                 815

tct tac ggg cat gta gat gta gca gct cta cta ata aag tat aat gca    2497
Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys Tyr Asn Ala
            820                 825                 830

tgt gtc aat gcc acg gac aaa tgg gct ttc aca cct ttg cac gaa gca    2545
Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala
        835                 840                 845

gcc caa aag gga cga aca cag ctt tgt gct ttg ttg cta gcc cat gga    2593
Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly
    850                 855                 860

gct gac ccg act ctt aaa aat cag gaa gga caa aca cct tta gat tta    2641
Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu
865                 870                 875                 880

gtt tca gca gat gat gtc agc gct ctt ctg aca gca gcc atg ccc cca    2689
Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala Met Pro Pro
                885                 890                 895

tct gct ctg ccc tct tgt tac aag cct caa gtg ctc aat ggt gtg aga    2737
Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn Gly Val Arg
            900                 905                 910

agc cca gga gcc act gca gat gct ctc tct tca ggt cca tct agc cca    2785
Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro Ser Ser Pro
        915                 920                 925

tca agc ctt tct gca gcc agc agt ctt gac aac tta tct ggg agt ttt    2833
Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser Gly Ser Phe
    930                 935                 940

tca gaa ctg tct tca gta gtt agt tca agt gga aca gag ggt gct tcc    2881
Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu Gly Ala Ser
945                 950                 955                 960

agt ttg gag aaa aag gag gtt cca gga gta gat ttt agc ata act caa    2929
Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser Ile Thr Gln
                965                 970                 975

ttc gta agg aat ctt gga ctt gag cac cta atg gat ata ttt gag aga    2977
Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile Phe Glu Arg
            980                 985                 990

gaa cag atc act ttg gat gta tta gtt gag atg ggg cac aag gag ctg    3025
Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His Lys Glu Leu
        995                1000                1005

aag gag att gga atc aat gct tat gga cat agg cac aaa cta att aaa    3073
Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys Leu Ile Lys
   1010                1015                1020

gga gtc gag aga ctt atc tcc gga caa caa ggt ctt aac cca tat tta    3121
Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn Pro Tyr Leu
1025                1030                1035                1040
```

```
act ttg aac acc tct ggt agt gga aca att ctt ata gat ctg tct cct    3169
Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp Leu Ser Pro
            1045                1050                1055

gat gat aaa gag ttt cag tct gtg gag gaa gag atg caa agt aca gtt    3217
Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln Ser Thr Val
        1060                1065                1070

cga gag cac aga gat gga ggt cat gca ggt gga atc ttc aac aga tac    3265
Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe Asn Arg Tyr
    1075                1080                1085

aat att ctc aag att cag aag gtt tgt aac aag aaa cta tgg gaa aga    3313
Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu Trp Glu Arg
1090                1095                1100

tac act cac cgg aga aaa gaa gtt tct gaa gaa aac cac aac cat gcc    3361
Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn His Ala
1105                1110                1115                1120

aat gaa cga atg cta ttt cat ggg tct cct ttt gtg aat gca att atc    3409
Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile
            1125                1130                1135

cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt ggt atg ttt gga    3457
His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly
        1140                1145                1150

gct ggc att tat ttt gct gaa aac tct tcc aaa agc aat caa tat gta    3505
Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val
    1155                1160                1165

tat gga att gga gga ggt act ggg tgt cca gtt cac aaa gac aga tct    3553
Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser
1170                1175                1180

tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg gta acc ttg gga    3601
Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly
1185                1190                1195                1200

aag tct ttc ctg cag ttc agt gca atg aaa atg gca cat tct cct cca    3649
Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro
            1205                1210                1215

ggt cat cac tca gtc act ggt agg ccc agt gta aat ggc cta gca tta    3697
Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu
        1220                1225                1230

gct gaa tat gtt att tac aga gga gaa cag gct tat cct gag tat tta    3745
Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu Tyr Leu
    1235                1240                1245

att act tac cag att atg agg cct gaa ggt atg gtc gat gga            3787
Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
1250                1255                1260

taaatagtta ttttaagaaa ctaattccac tgaacctaaa atcatcaaag cagcagtggc 3847

ctctacgttt tactcctttg ctgaaaaaaa atcatcttgc ccacaggcct gtggcaaaag 3907

gataaaaatg tgaacgaagt ttaacattct gacttgataa agctttaata atgtacagtg 3967

ttttctaaat atttcctgtt ttttcagcac tttaacagat gccattccag gttaaactgg 4027

gttgtctgta ctaaattata aacagagtta acttgaacct tttatatgtt atgcattgat 4087

tctaacaaac tgtaatgccc tcaacagaac taattttact aatacaatac tgtgttcttt 4147

aaaacacagc atttacactg aatacaattt catttgtaaa actgtaaata agagcttttg 4207
```

tactagccca gtatttattt acattgcttt gtaatataaa tctgttttag aactgcagcg 4267

gtttacaaaa ttttttcata tgtattgttc atctatactt catcttacat cgtcatgatt 4327

gagtgatctt tacatttgat tccagaggct atgttcagtt gttagttggg aaagattgag 4387

ttatcagatt taatttgcc 4406

<210> 107
<211> 1262
<212> PRT
<213> Homo sapiens

<400> 107
Glu Leu Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln
1 5 10 15

Asp Pro Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg
20 25 30

Gly Ala Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr
35 40 45

Ala Pro Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser
50 55 60

Ala Ser Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu
65 70 75 80

Leu Leu Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile
85 90 95

Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala
100 105 110

Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
115 120 125

Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys
130 135 140

Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe
145 150 155 160

Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn
165 170 175

Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His
180 185 190

Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu Leu Leu Arg
195 200 205

His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu
210 215 220

His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu
225 230 235 240

Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly Arg Thr Ala
245 250 255

Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr
260 265 270

Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn Glu Glu Lys
275 280 285

Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp
290 295 300

Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val
305 310 315 320

Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val His Ala Lys
325 330 335

Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His
340 345 350

Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys Val Asn Ala
355 360 365

Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn
370 375 380

Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr
385 390 395 400

Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro
405 410 415

Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu
420 425 430

Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser
435 440 445

Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu Thr Ala Leu
450 455 460

His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu
465 470 475 480

Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe
485 490 495

Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn Asp Val Val
500 505 510

Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu
515 520 525

Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr
530 535 540

Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu
545 550 555 560

Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu
565 570 575

Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu
580 585 590

Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys
595 600 605

Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr
610 615 620

Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr
625 630 635 640

Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu
645 650 655

Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu
660 665 670

Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys
675 680 685

Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys
690 695 700

Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp
705 710 715 720

Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp Ile Gln
725 730 735

Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly
740 745 750

Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn Val Asn Cys
755 760 765

Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu Ala Ala Gly
770 775 780

Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His Gly Ala Asp
785 790 795 800

Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala
805 810 815

Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys Tyr Asn Ala
820 825 830

Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala
835 840 845

Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly
850 855 860

Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu
865 870 875 880

Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala Met Pro Pro
885 890 895

Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn Gly Val Arg
900 905 910

Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro Ser Ser Pro
915 920 925

Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser Gly Ser Phe
930 935 940

Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu Gly Ala Ser
945 950 955 960

Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser Ile Thr Gln
965 970 975

Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile Phe Glu Arg
980 985 990

Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His Lys Glu Leu
995 1000 1005

Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys Leu Ile Lys
1010 1015 1020

Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn Pro Tyr Leu
1025 1030 1035 1040

Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp Leu Ser Pro
1045 1050 1055

Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln Ser Thr Val
1060 1065 1070

Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe Asn Arg Tyr
1075 1080 1085

Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu Trp Glu Arg
1090 1095 1100

Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn His Ala
1105 1110 1115 1120

Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile
1125 1130 1135

His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly
1140 1145 1150

Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val
1155 1160 1165

Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser
1170 1175 1180

Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly
1185 1190 1195 1200

Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro
1205 1210 1215

Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu
1220 1225 1230

Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu Tyr Leu
1235 1240 1245

Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
1250 1255 1260

<210> 108
<211> 436
<212> DNA
<213> Homo sapiens

<400> 108
tttttttgc agttctaaaa cagatttata ttacaaagca atgtaaataa atactgggct 60
agtacaaaag ctcttattta cagttttaca aatgaaattg tattcagtgt aaatgctgtg 120
ttttaaagaa cacagtattg tattagtaaa attagttctg ttgagggcat tacagtttgt 180
tagaatcaat gcataacata taaaaggttc aagttaactc tgtttataat ttagtacaga 240
caacccagtt taacctggga tgggcatctg ttaaagtgct ggaaaaaaca gggaaatatt 300
taggaaaaca ctggtacatt atttaaaggc tttntccaag gtcaggantg tttaaacttc 360
gtttcacatt tttatccntt tggccacggc ctgtggggcn aggatggatt ttttttccgg 420
ccaagggtgt taaacg 436

<210> 109
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 109
cgcctgagaa ggtgaacagc c 21

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 110
acgcctcgaa cagctctcgg 20

<210> 111
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 111
gcgtgggcgc ggccatggga ctg 23

<210> 112
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 112
cagcgcgaat ccgccgtccg 20

<210> 113
<211> 620
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(620)

<400> 113
tt aaa aca aca aca aca aaa aac aca ata tgc agg atc gtt cgg ctt 47
Lys Thr Thr Thr Thr Lys Asn Thr Ile Cys Arg Ile Val Arg Leu
1 5 10 15

cag cag aac cca ccg caa aga tgg cgg tgg gac gaa gcc cct tct ccc 95
Gln Gln Asn Pro Pro Gln Arg Trp Arg Trp Asp Glu Ala Pro Ser Pro
20 25 30

gcc gcc gaa gcc tct cgc ctc aca ttt ccc aca aac cct tcg cgc cgc 143
Ala Ala Glu Ala Ser Arg Leu Thr Phe Pro Thr Asn Pro Ser Arg Arg
35 40 45

ctc gct agc cga aac ctg ccc agc cgg tgc ccg gcc act gcg cac gcg 191
Leu Ala Ser Arg Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala
50 55 60

cgg gac gac gtc acg tgc gct ccc ggg gct gga cgg agc tgg cag gag 239
Arg Asp Asp Val Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu
65 70 75

ctg gca gga ggg gcc ttg cca gct tcc gcc gcc gcg tcg ttt cag gac 287
Leu Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp
80 85 90 95

ccg gac ggc gga ttc gcg ctg cct ccg ccg ccg cgg ggc agc cgg ggg 335
Pro Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly
100 105 110

gca ggg agc cca gcg agg ggc gcg cgt ggg cgc ggc cat ggg act gcg 383
Ala Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala
115 120 125

ccg gat ccg gtg aca gca ggg agc caa gcg gcc cgg gcc ctg agc gcg 431
Pro Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala
130 135 140

tct tct ccg ggg ggc ctc gcc ctc ctg ctc gcg ggg ccg ggg ctc ctg 479
Ser Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu
145 150 155

ctc cgg ttg ctg gcg ctg ttg ctg gct gtg gcg gcg gcc agg atc atg 527
Leu Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met
160 165 170 175

tcg ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc gcg agc gcc gcg 575
Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala
180 185 190

gcc gag gcc gtg gag ccg gcc gcc cga gag ctg ttc gag gcg tgc 620
Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
195 200 205

<210> 114
<211> 206
<212> PRT
<213> Homo sapiens

<400> 114
Lys Thr Thr Thr Thr Lys Asn Thr Ile Cys Arg Ile Val Arg Leu Gln
1 5 10 15

Gln Asn Pro Pro Gln Arg Trp Arg Trp Asp Glu Ala Pro Ser Pro Ala
20 25 30

Ala Glu Ala Ser Arg Leu Thr Phe Pro Thr Asn Pro Ser Arg Arg Leu
35 40 45

Ala Ser Arg Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala Arg
50 55 60

Asp Asp Val Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu
65 70 75 80

Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro
85 90 95

Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala
100 105 110

Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala Pro
115 120 125

Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser
130 135 140

Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu
145 150 155 160

```
Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met Ser
                165                 170                 175

Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala
            180                 185                 190

Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
        195                 200                 205


<210> 115
<211> 1039
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (287)..(1039)

<400> 115
gtacaatatt gatttacaaa aagttcctct aatcaatcct gagctaataa cttactgtgg 60

aaagagtaat tgatcagagc catccctcca attggagtca actttcatga ctgttcggat 120

ttcctttatt ttgggggcag ttcatccaaa cttctattaa acggcaacta gttcactttt 180

gagaagtggt ttacaagaaa caacaacaac aacaacaaag cagttgcgga ggaaagaaaa 240

gagacaaagt aaaaaaaacg gaaaagaaat ctcccaggag aaaggg atg tgg aag    295
                                                   Met Trp Lys
                                                   1

ctg aaa aca cgg aca att tcc aca gta aga ctt cca aaa gaa tgt gca    343
Leu Lys Thr Arg Thr Ile Ser Thr Val Arg Leu Pro Lys Glu Cys Ala
    5                   10                  15

aga tcc gag caa aac ttt caa ggg ctc ttt ttc agt gta atg gta gtg    391
Arg Ser Glu Gln Asn Phe Gln Gly Leu Phe Phe Ser Val Met Val Val
20                  25                  30                  35

aga aag ttc agc ctg gaa agc cca ggg ctt aaa aca aca aca aca aaa    439
Arg Lys Phe Ser Leu Glu Ser Pro Gly Leu Lys Thr Thr Thr Thr Lys
                40                  45                  50

aac aca ata tgc agg atc gtt cgg ctt cag cag aac cca ccg caa aga    487
Asn Thr Ile Cys Arg Ile Val Arg Leu Gln Gln Asn Pro Pro Gln Arg
            55                  60                  65

tgg cgg tgg gac gaa gcc cct tct ccc gcc gcc gaa gcc tct cgc ctc    535
Trp Arg Trp Asp Glu Ala Pro Ser Pro Ala Ala Glu Ala Ser Arg Leu
        70                  75                  80

aca ttt ccc aca aac cct tcg cgc cgc ctc gct agc cga aac ctg ccc    583
Thr Phe Pro Thr Asn Pro Ser Arg Arg Leu Ala Ser Arg Asn Leu Pro
    85                  90                  95

agc cgg tgc ccg gcc act gcg cac gcg cgg gac gac gtc acg tgc gct    631
Ser Arg Cys Pro Ala Thr Ala His Ala Arg Asp Asp Val Thr Cys Ala
100                 105                 110                 115

ccc ggg gct gga cgg agc tgg cag gag ctg gca gga ggg gcc ttg cca    679
Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu Ala Gly Gly Ala Leu Pro
                120                 125                 130

gct tcc gcc gcc gcg tcg ttt cag gac ccg gac ggc gga ttc gcg ctg    727
Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro Asp Gly Gly Phe Ala Leu
            135                 140                 145
```

```
cct ccg ccg ccg cgg ggc agc cgg ggg gca ggg agc cca gcg agg ggc      775
Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala Gly Ser Pro Ala Arg Gly
        150                 155                 160

gcg cgt ggg cgc ggc cat ggg act gcg ccg gat ccg gtg aca gca ggg      823
Ala Arg Gly Arg Gly His Gly Thr Ala Pro Asp Pro Val Thr Ala Gly
    165                 170                 175

agc caa gcg gcc cgg gcc ctg agc gcg tct tct ccg ggg ggc ctc gcc      871
Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser Ser Pro Gly Gly Leu Ala
180                 185                 190                 195

ctc ctg ctc gcg ggg ccg ggg ctc ctg ctc cgg ttg ctg gcg ctg ttg      919
Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu Arg Leu Leu Ala Leu Leu
                200                 205                 210

ctg gct gtg gcg gcg gcc agg atc atg tcg ggt cgc cgc tgc gcc ggc      967
Leu Ala Val Ala Ala Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly
            215                 220                 225

ggg gga gcg gcc tgc gcg agc gcc gcg gcc gag gcc gtg gag ccg gcc      1015
Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala
        230                 235                 240

gcc cga gag ctg ttc gag gcg tgc                                      1039
Ala Arg Glu Leu Phe Glu Ala Cys
    245                 250


<210> 116
<211> 251
<212> PRT
<213> Homo sapiens

<400> 116
Met Trp Lys Leu Lys Thr Arg Thr Ile Ser Thr Val Arg Leu Pro Lys
 1               5                  10                  15

Glu Cys Ala Arg Ser Glu Gln Asn Phe Gln Gly Leu Phe Phe Ser Val
            20                  25                  30

Met Val Val Arg Lys Phe Ser Leu Glu Ser Pro Gly Leu Lys Thr Thr
        35                  40                  45

Thr Thr Lys Asn Thr Ile Cys Arg Ile Val Arg Leu Gln Gln Asn Pro
    50                  55                  60

Pro Gln Arg Trp Arg Trp Asp Glu Ala Pro Ser Pro Ala Ala Glu Ala
65                  70                  75                  80

Ser Arg Leu Thr Phe Pro Thr Asn Pro Ser Arg Arg Leu Ala Ser Arg
                85                  90                  95

Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala Arg Asp Asp Val
            100                 105                 110

Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu Ala Gly Gly
        115                 120                 125

Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro Asp Gly Gly
    130                 135                 140

Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala Gly Ser Pro
145                 150                 155                 160

Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala Pro Asp Pro Val
                165                 170                 175

Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser Ser Pro Gly
            180                 185                 190
```

```
Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu Arg Leu Leu
        195                 200                 205

Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met Ser Gly Arg Arg
    210                 215                 220

Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu Ala Val
225                 230                 235                 240

Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
                245                 250


<210> 117
<211> 473
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(473)

<400> 117
ct agc cga aac ctg ccc agc cgg tgc ccg gcc act gcg cac gcg cgg      47
   Ser Arg Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala Arg
     1               5                  10                  15

gac gac gtc acg tgc gct ccc ggg gct gga cgg agc tgg cag gag ctg     95
Asp Asp Val Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu
                20                  25                  30

gca gga ggg gcc ttg cca gct tcc gcc gcc gcg tcg ttt cag gac ccg     143
Ala Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro
            35                  40                  45

gac ggc gga ttc gcg ctg cct ccg ccg ccg cgg ggc agc cgg ggg gca     191
Asp Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala
        50                  55                  60

ggg agc cca gcg agg ggc gcg cgt ggg cgc ggc cat ggg act gcg ccg     239
Gly Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala Pro
    65                  70                  75

gat ccg gtg aca gca ggg agc caa gcg gcc cgg gcc ctg agc gcg tct     287
Asp Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser
80                  85                  90                  95

tct ccg ggg ggc ctc gcc ctc ctg ctc gcg ggg ccg ggg ctc ctg ctc     335
Ser Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu
                100                 105                 110

cgg ttg ctg gcg ctg ttg ctg gct gtg gcg gcg gcc agg atc atg tcg     383
Arg Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met Ser
            115                 120                 125

ggt cgc cgc tgc gcc ggc ggg gga gcg gcc tgc gcg agc gcc gcg gcc     431
Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala
        130                 135                 140

gag gcc gtg gag ccg gcc gcc cga gag ctg ttc gag gcg tgc             473
Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
    145                 150                 155

<210> 118
<211> 157
<212> PRT
<213> Homo sapiens
```

<400> 118
Ser Arg Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala Arg Asp
1 5 10 15

Asp Val Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu Ala
20 25 30

Gly Gly Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro Asp
35 40 45

Gly Gly Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala Gly
50 55 60

Ser Pro Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala Pro Asp
65 70 75 80

Pro Val Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser Ser
85 90 95

Pro Gly Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu Arg
100 105 110

Leu Leu Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met Ser Gly
115 120 125

Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu
130 135 140

Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
145 150 155

<210> 119
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 119
gttcctctaa tcaatcctga gc 22

<210> 120
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 120
ggaaagagta attgatcaga gccatc 26

<210> 121
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 121
cgccgaagcc tctcgcctca catttcc 27

<210> 122

<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 122
ggaaatgtga ggcgagaggc ttcggcg 27

<210> 123
<211> 659
<212> DNA
<213> Homo sapiens

<400> 123
ggaaagagta attgatcaga gccatccctc caattggagt caacttccat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga ttcgcgctg 659

<210> 124
<211> 669
<212> DNA
<213> Homo sapiens

<400> 124
ggaaagagta attgatcaga gccatccctc caattggagt caactttcat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
gggccaggag gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga 660
ttcgcgctg 669

<210> 125
<211> 659
<212> DNA
<213> Homo sapiens

<400> 125
ggaaagagta attgatcaga gccatccctc caattggagt caacttccat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga ttcgcgctg 659

<210> 126

<211> 659
<212> DNA
<213> Homo sapiens

<400> 126
ggaaagagta attgatcaga gccatccctc caattggagt caacttccat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga ttcgcgctg 659

<210> 127
<211> 659
<212> DNA
<213> Homo sapiens

<400> 127
ggaaagagta attgatcaga gccatccctc caattggagt caacttccat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga ttcgcgctg 659

<210> 128
<211> 669
<212> DNA
<213> Homo sapiens

<400> 128
ggaaagagta attgatcaga gccatccctc caattggagt caactttcat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcatg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600
ctggcaggag gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga 660
ttcgcgctg 669

<210> 129
<211> 597
<212> DNA
<213> Homo sapiens

<400> 129
ggaaagagta attgatcaga gccatccctc caattggagt caactttcat gactgttcgg 60
atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120
ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180
aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240
aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300
tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360
cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420

```
ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480
tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540
actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcag    597


<210> 130
<211> 10
<212> DNA
<213> Homo sapiens

<400> 130
gagctggcag                                                          10


<210> 131
<211> 30
<212> DNA
<213> Homo sapiens

<400> 131
gggctggacg gagctggcag gaggggcctt                                    30


<210> 132
<211> 5002
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (229)..(4383)

<400> 132
ggaaagagta attgatcaga gccatccctc caattggagt caactttcat gactgttcgg 60

atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120

ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180

aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaaggg atg tgg aag  237
                                                     Met Trp Lys
                                                         1

ctg aaa aca cgg aca att tcc aca gta aga ctt cca aaa gaa tgt gca   285
Leu Lys Thr Arg Thr Ile Ser Thr Val Arg Leu Pro Lys Glu Cys Ala
    5                   10                  15

aga tcc gag caa aac ttt caa ggg ctc ttt ttc agt gta atg gta gtg   333
Arg Ser Glu Gln Asn Phe Gln Gly Leu Phe Phe Ser Val Met Val Val
20                  25                  30                  35

aga aag ttc agc ctg gaa agc cca ggg ctt aaa aca aca aca aca aaa   381
Arg Lys Phe Ser Leu Glu Ser Pro Gly Leu Lys Thr Thr Thr Thr Lys
                40                  45                  50

aac aca ata tgc agg atc gtt cgg ctt cag cag aac cca ccg caa aga   429
Asn Thr Ile Cys Arg Ile Val Arg Leu Gln Gln Asn Pro Pro Gln Arg
            55                  60                  65

tgg cgg tgg gac gaa gcc cct tct ccc gcc gcc gaa gcc tct cgc ctc   477
Trp Arg Trp Asp Glu Ala Pro Ser Pro Ala Ala Glu Ala Ser Arg Leu
        70                  75                  80

aca ttt ccc aca aac cct tcg cgc cgc ctc gct agc cga aac ctg ccc   525
Thr Phe Pro Thr Asn Pro Ser Arg Arg Leu Ala Ser Arg Asn Leu Pro
    85                  90                  95

agc cgg tgc ccg gcc act gcg cac gcg cgg gac gac gtc acg tgc gct   573
Ser Arg Cys Pro Ala Thr Ala His Ala Arg Asp Asp Val Thr Cys Ala
100                 105                 110                 115
```

```
ccc ggg gct gga cgg agc tgg cag gag ctg gca gga ggg gcc ttg cca     621
Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu Ala Gly Gly Ala Leu Pro
            120                 125                 130

gct tcc gcc gcc gcg tcg ttt cag gac ccg gac ggc gga ttc gcg ctg     669
Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro Asp Gly Gly Phe Ala Leu
        135                 140                 145

cct ccg ccg ccg cgg ggc agc cgg ggg gca ggg agc cca gcg agg ggc     717
Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala Gly Ser Pro Ala Arg Gly
    150                 155                 160

gcg cgt ggg cgc ggc cat ggg act gcg ccg gat ccg gtg aca gca ggg     765
Ala Arg Gly Arg Gly His Gly Thr Ala Pro Asp Pro Val Thr Ala Gly
    165                 170                 175

agc caa gcg gcc cgg gcc ctg agc gcg tct tct ccg ggg ggc ctc gcc     813
Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser Ser Pro Gly Gly Leu Ala
180                 185                 190                 195

ctc ctg ctc gcg ggg ccg ggg ctc ctg ctc cgg ttg ctg gcg ctg ttg     861
Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu Arg Leu Leu Ala Leu Leu
            200                 205                 210

ctg gct gtg gcg gcg gcc agg atc atg tcg ggt cgc cgc tgc gcc ggc     909
Leu Ala Val Ala Ala Ala Arg Ile Met Ser Gly Arg Arg Cys Ala Gly
        215                 220                 225

ggg gga gcg gcc tgc gcg agc gcc gcg gcc gag gcc gtg gag ccg gcc     957
Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu Ala Val Glu Pro Ala
    230                 235                 240

gcc cga gag ctg ttc gag gcg tgc cgc aac ggg gac gtg gaa cga gtc     1005
Ala Arg Glu Leu Phe Glu Ala Cys Arg Asn Gly Asp Val Glu Arg Val
    245                 250                 255

aag agg ctg gtg acg cct gag aag gtg aac agc cgc gac acg gcg ggc     1053
Lys Arg Leu Val Thr Pro Glu Lys Val Asn Ser Arg Asp Thr Ala Gly
260                 265                 270                 275

agg aaa tcc acc ccg ctg cac ttc gcc gca ggt ttt ggg cgg aaa gac     1101
Arg Lys Ser Thr Pro Leu His Phe Ala Ala Gly Phe Gly Arg Lys Asp
            280                 285                 290

gta gtt gaa tat ttg ctt cag aat ggt gca aat gtc caa gca cgt gat     1149
Val Val Glu Tyr Leu Leu Gln Asn Gly Ala Asn Val Gln Ala Arg Asp
        295                 300                 305

gat ggg ggc ctt att cct ctt cat aat gca tgc tct ttt ggt cat gct     1197
Asp Gly Gly Leu Ile Pro Leu His Asn Ala Cys Ser Phe Gly His Ala
    310                 315                 320

gaa gta gtc aat ctc ctt ttg cga cat ggt gca gac ccc aat gct cga     1245
Glu Val Val Asn Leu Leu Leu Arg His Gly Ala Asp Pro Asn Ala Arg
    325                 330                 335

gat aat tgg aat tat act cct ctc cat gaa gct gca att aaa gga aag     1293
Asp Asn Trp Asn Tyr Thr Pro Leu His Glu Ala Ala Ile Lys Gly Lys
340                 345                 350                 355

att gat gtt tgc att gtg ctg tta cag cat gga gct gag cca acc atc     1341
Ile Asp Val Cys Ile Val Leu Leu Gln His Gly Ala Glu Pro Thr Ile
            360                 365                 370

cga aat aca gat gga agg aca gca ttg gat tta gca gat cca tct gcc     1389
Arg Asn Thr Asp Gly Arg Thr Ala Leu Asp Leu Ala Asp Pro Ser Ala
        375                 380                 385
```

```
aaa gca gtg ctt act ggt gaa tat aag aaa gat gaa ctc tta gaa agt     1437
Lys Ala Val Leu Thr Gly Glu Tyr Lys Lys Asp Glu Leu Leu Glu Ser
        390                 395                 400

gcc agg agt ggc aat gaa gaa aaa atg atg gct cta ctc aca cca tta     1485
Ala Arg Ser Gly Asn Glu Glu Lys Met Met Ala Leu Leu Thr Pro Leu
    405                 410                 415

aat gtc aac tgc cac gca agt gat ggc aga aag tca act cca tta cat     1533
Asn Val Asn Cys His Ala Ser Asp Gly Arg Lys Ser Thr Pro Leu His
420                 425                 430                 435

ttg gca gca gga tat aac aga gta aag att gta cag ctg tta ctg caa     1581
Leu Ala Ala Gly Tyr Asn Arg Val Lys Ile Val Gln Leu Leu Leu Gln
            440                 445                 450

cat gga gct gat gtc cat gct aaa gat aaa ggt gat ctg gta cca tta     1629
His Gly Ala Asp Val His Ala Lys Asp Lys Gly Asp Leu Val Pro Leu
            455                 460                 465

cac aat gcc tgt tct tat ggt cat tat gaa gta act gaa ctt ttg gtc     1677
His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val
        470                 475                 480

aag cat ggt gcc tgt gta aat gca atg gac ttg tgg caa ttc act cct     1725
Lys His Gly Ala Cys Val Asn Ala Met Asp Leu Trp Gln Phe Thr Pro
    485                 490                 495

ctt cat gag gca gct tct aag aac agg gtt gaa gta tgt tct ctt ctc     1773
Leu His Glu Ala Ala Ser Lys Asn Arg Val Glu Val Cys Ser Leu Leu
500                 505                 510                 515

tta agt tat ggt gca gac cca aca ctg ctc aat tgt cac aat aaa agt     1821
Leu Ser Tyr Gly Ala Asp Pro Thr Leu Leu Asn Cys His Asn Lys Ser
            520                 525                 530

gct ata gac ttg gct ccc aca cca cag tta aaa gaa aga tta gca tat     1869
Ala Ile Asp Leu Ala Pro Thr Pro Gln Leu Lys Glu Arg Leu Ala Tyr
            535                 540                 545

gaa ttt aaa ggc cac tcg ttg ctg caa gct gca cga gaa gct gat gtt     1917
Glu Phe Lys Gly His Ser Leu Leu Gln Ala Ala Arg Glu Ala Asp Val
        550                 555                 560

act cga atc aaa aaa cat ctc tct ctg gaa atg gtg aat ttc aag cat     1965
Thr Arg Ile Lys Lys His Leu Ser Leu Glu Met Val Asn Phe Lys His
    565                 570                 575

cct caa aca cat gaa aca gca ttg cat tgt gct gct gca tct cca tat     2013
Pro Gln Thr His Glu Thr Ala Leu His Cys Ala Ala Ala Ser Pro Tyr
580                 585                 590                 595

ccc aaa aga aag caa ata tgt gaa ctg ttg cta aga aaa gga gca aac     2061
Pro Lys Arg Lys Gln Ile Cys Glu Leu Leu Leu Arg Lys Gly Ala Asn
            600                 605                 610

atc aat gaa aag act aaa gaa ttc ttg act cct ctg cac gtg gca tct     2109
Ile Asn Glu Lys Thr Lys Glu Phe Leu Thr Pro Leu His Val Ala Ser
            615                 620                 625

gag aaa gct cat aat gat gtt gtt gaa gta gtg gtg aaa cat gaa gca     2157
Glu Lys Ala His Asn Asp Val Val Glu Val Val Val Lys His Glu Ala
        630                 635                 640

aag gtt aat gct ctg gat aat ctt ggt cag act tct cta cac aga gct     2205
Lys Val Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu His Arg Ala
    645                 650                 655
```

gca tat tgt ggt cat cta caa acc tgc cgc cta ctc ctg agc tat ggg 2253
Ala Tyr Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu Ser Tyr Gly
660 665 670 675

tgt gat cct aac att ata tcc ctt cag ggc ttt act gct tta cag atg 2301
Cys Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala Leu Gln Met
680 685 690

gga aat gaa aat gta cag caa ctc ctc caa gag ggt atc tca tta ggt 2349
Gly Asn Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile Ser Leu Gly
695 700 705

aat tca gag gca gac aga caa ttg ctg gaa gct gca aag gct gga gat 2397
Asn Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys Ala Gly Asp
710 715 720

gtc gaa act gta aaa aaa ctg tgt act gtt cag agt gtc aac tgc aga 2445
Val Glu Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val Asn Cys Arg
725 730 735

gac att gaa ggg cgt cag tct aca cca ctt cat ttt gca gct ggg tat 2493
Asp Ile Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala Ala Gly Tyr
740 745 750 755

aac aga gtg tcc gtg gtg gaa tat ctg cta cag cat gga gct gat gtg 2541
Asn Arg Val Ser Val Val Glu Tyr Leu Leu Gln His Gly Ala Asp Val
760 765 770

cat gct aaa gat aaa gga ggc ctt gta cct ttg cac aat gca tgt tct 2589
His Ala Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser
775 780 785

tat gga cat tat gaa gtt gca gaa ctt ctt gtt aaa cat gga gca gta 2637
Tyr Gly His Tyr Glu Val Ala Glu Leu Leu Val Lys His Gly Ala Val
790 795 800

gtt aat gta gct gat tta tgg aaa ttt aca cct tta cat gaa gca gca 2685
Val Asn Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His Glu Ala Ala
805 810 815

gca aaa gga aaa tat gaa att tgc aaa ctt ctg ctc cag cat ggt gca 2733
Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln His Gly Ala
820 825 830 835

gac cct aca aaa aaa aac agg gat gga aat act cct ttg gat ctt gtt 2781
Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu Asp Leu Val
840 845 850

aaa gat gga gat aca gat att caa gat ctg ctt agg gga gat gca gct 2829
Lys Asp Gly Asp Thr Asp Ile Gln Asp Leu Leu Arg Gly Asp Ala Ala
855 860 865

ttg cta gat gct gcc aag aag ggt tgt tta gcc aga gtg aag aag ttg 2877
Leu Leu Asp Ala Ala Lys Lys Gly Cys Leu Ala Arg Val Lys Lys Leu
870 875 880

tct tct cct gat aat gta aat tgc cgc gat acc caa ggc aga cat tca 2925
Ser Ser Pro Asp Asn Val Asn Cys Arg Asp Thr Gln Gly Arg His Ser
885 890 895

aca cct tta cat tta gca gct ggt tat aat aat tta gaa gtt gca gag 2973
Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Leu Glu Val Ala Glu
900 905 910 915

tat ttg tta caa cac gga gct gat gtg aat gcc caa gac aaa gga gga 3021
Tyr Leu Leu Gln His Gly Ala Asp Val Asn Ala Gln Asp Lys Gly Gly
920 925 930

```
ctt att cct tta cat aat gca gca tct tac ggg cat gta gat gta gca     3069
Leu Ile Pro Leu His Asn Ala Ala Ser Tyr Gly His Val Asp Val Ala
            935                 940                 945

gct cta cta ata aag tat aat gca tgt gtc aat gcc acg gac aaa tgg     3117
Ala Leu Leu Ile Lys Tyr Asn Ala Cys Val Asn Ala Thr Asp Lys Trp
        950                 955                 960

gct ttc aca cct ttg cac gaa gca gcc caa aag gga cga aca cag ctt     3165
Ala Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr Gln Leu
    965                 970                 975

tgt gct ttg ttg cta gcc cat gga gct gac ccg act ctt aaa aat cag     3213
Cys Ala Leu Leu Leu Ala His Gly Ala Asp Pro Thr Leu Lys Asn Gln
980                 985                 990                 995

gaa gga caa aca cct tta gat tta gtt tca gca gat gat gtc agc gct     3261
Glu Gly Gln Thr Pro Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala
                1000                1005                1010

ctt ctg aca gca gcc atg ccc cca tct gct ctg ccc tct tgt tac aag     3309
Leu Leu Thr Ala Ala Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys
            1015                1020                1025

cct caa gtg ctc aat ggt gtg aga agc cca gga gcc act gca gat gct     3357
Pro Gln Val Leu Asn Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala
        1030                1035                1040

ctc tct tca ggt cca tct agc cca tca agc ctt tct gca gcc agc agt     3405
Leu Ser Ser Gly Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser
    1045                1050                1055

ctt gac aac tta tct ggg agt ttt tca gaa ctg tct tca gta gtt agt     3453
Leu Asp Asn Leu Ser Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser
1060                1065                1070                1075

tca agt gga aca gag ggt gct tcc agt ttg gag aaa aag gag gtt cca     3501
Ser Ser Gly Thr Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro
                1080                1085                1090

gga gta gat ttt agc ata act caa ttc gta agg aat ctt gga ctt gag     3549
Gly Val Asp Phe Ser Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu
            1095                1100                1105

cac cta atg gat ata ttt gag aga gaa cag atc act ttg gat gta tta     3597
His Leu Met Asp Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu
        1110                1115                1120

gtt gag atg ggg cac aag gag ctg aag gag att gga atc aat gct tat     3645
Val Glu Met Gly His Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr
    1125                1130                1135

gga cat agg cac aaa cta att aaa gga gtc gag aga ctt atc tcc gga     3693
Gly His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly
1140                1145                1150                1155

caa caa ggt ctt aac cca tat tta act ttg aac acc tct ggt agt gga     3741
Gln Gln Gly Leu Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly
                1160                1165                1170

aca att ctt ata gat ctg tct cct gat gat aaa gag ttt cag tct gtg     3789
Thr Ile Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val
            1175                1180                1185

gag gaa gag atg caa agt aca gtt cga gag cac aga gat gga ggt cat     3837
Glu Glu Glu Met Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly His
        1190                1195                1200
```

gca ggt gga atc ttc aac aga tac aat att ctc aag att cag aag gtt 3885
Ala Gly Gly Ile Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val
1205 1210 1215

tgt aac aag aaa cta tgg gaa aga tac act cac cgg aga aaa gaa gtt 3933
Cys Asn Lys Lys Leu Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val
1220 1225 1230 1235

tct gaa gaa aac cac aac cat gcc aat gaa cga atg cta ttt cat ggg 3981
Ser Glu Glu Asn His Asn His Ala Asn Glu Arg Met Leu Phe His Gly
1240 1245 1250

tct cct ttt gtg aat gca att atc cac aaa ggc ttt gat gaa agg cat 4029
Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg His
1255 1260 1265

gcg tac ata ggt ggt atg ttt gga gct ggc att tat ttt gct gaa aac 4077
Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn
1270 1275 1280

tct tcc aaa agc aat caa tat gta tat gga att gga gga ggt act ggg 4125
Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly
1285 1290 1295

tgt cca gtt cac aaa gac aga tct tgt tac att tgc cac agg cag ctg 4173
Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu
1300 1305 1310 1315

ctc ttt tgc cgg gta acc ttg gga aag tct ttc ctg cag ttc agt gca 4221
Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala
1320 1325 1330

atg aaa atg gca cat tct cct cca ggt cat cac tca gtc act ggt agg 4269
Met Lys Met Ala His Ser Pro Pro Gly His His Ser Val Thr Gly Arg
1335 1340 1345

ccc agt gta aat ggc cta gca tta gct gaa tat gtt att tac aga gga 4317
Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly
1350 1355 1360

gaa cag gct tat cct gag tat tta att act tac cag att atg agg cct 4365
Glu Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro
1365 1370 1375

gaa ggt atg gtc gat gga taaatagtta ttttaagaaa ctaattccac 4413
Glu Gly Met Val Asp Gly
1380 1385

tgaacctaaa atcatcaaag cagcagtggc ctctacgttt tactcctttg ctgaaaaaaa 4473

atcatcttgc ccacaggcct gtggcaaaag gataaaaatg tgaacgaagt ttaacattct 4533

gacttgataa agctttaata atgtacagtg ttttctaaat atttcctgtt ttttcagcac 4593

tttaacagat gccattccag gttaaactgg gttgtctgta ctaaattata aacagagtta 4653

acttgaacct tttatatgtt atgcattgat tctaacaaac tgtaatgccc tcaacagaac 4713

taattttact aatacaatac tgtgttcttt aaaacacagc atttacactg aatacaattt 4773

catttgtaaa actgtaaata agagcttttg tactagccca gtatttattt acattgcttt 4833

gtaatataaa tctgttttag aactgcagcg gtttacaaaa ttttttcata tgtattgttc 4893

atctatactt catcttacat cgtcatgatt gagtgatctt tacatttgat tccagaggct 4953

atgttcagtt gttagttggg aaagattgag ttatcagatt taatttgcc 5002

<210> 133
<211> 1385
<212> PRT
<213> Homo sapiens

<400> 133
Met Trp Lys Leu Lys Thr Arg Thr Ile Ser Thr Val Arg Leu Pro Lys
1 5 10 15

Glu Cys Ala Arg Ser Glu Gln Asn Phe Gln Gly Leu Phe Phe Ser Val
20 25 30

Met Val Val Arg Lys Phe Ser Leu Glu Ser Pro Gly Leu Lys Thr Thr
35 40 45

Thr Thr Lys Asn Thr Ile Cys Arg Ile Val Arg Leu Gln Gln Asn Pro
50 55 60

Pro Gln Arg Trp Arg Trp Asp Glu Ala Pro Ser Pro Ala Ala Glu Ala
65 70 75 80

Ser Arg Leu Thr Phe Pro Thr Asn Pro Ser Arg Arg Leu Ala Ser Arg
85 90 95

Asn Leu Pro Ser Arg Cys Pro Ala Thr Ala His Ala Arg Asp Asp Val
100 105 110

Thr Cys Ala Pro Gly Ala Gly Arg Ser Trp Gln Glu Leu Ala Gly Gly
115 120 125

Ala Leu Pro Ala Ser Ala Ala Ala Ser Phe Gln Asp Pro Asp Gly Gly
130 135 140

Phe Ala Leu Pro Pro Pro Pro Arg Gly Ser Arg Gly Ala Gly Ser Pro
145 150 155 160

Ala Arg Gly Ala Arg Gly Arg Gly His Gly Thr Ala Pro Asp Pro Val
165 170 175

Thr Ala Gly Ser Gln Ala Ala Arg Ala Leu Ser Ala Ser Ser Pro Gly
180 185 190

Gly Leu Ala Leu Leu Leu Ala Gly Pro Gly Leu Leu Leu Arg Leu Leu
195 200 205

Ala Leu Leu Leu Ala Val Ala Ala Ala Arg Ile Met Ser Gly Arg Arg
210 215 220

Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu Ala Val
225 230 235 240

Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys Arg Asn Gly Asp Val
245 250 255

Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys Val Asn Ser Arg Asp
260 265 270

Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe Ala Ala Gly Phe Gly
275 280 285

Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn Gly Ala Asn Val Gln
290 295 300

Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His Asn Ala Cys Ser Phe
305 310 315 320

Gly His Ala Glu Val Val Asn Leu Leu Leu Arg His Gly Ala Asp Pro
325 330 335

Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu His Glu Ala Ala Ile
340 345 350

Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu Gln His Gly Ala Glu
355 360 365

Pro Thr Ile Arg Asn Thr Asp Gly Arg Thr Ala Leu Asp Leu Ala Asp
370 375 380

Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr Lys Lys Asp Glu Leu
385 390 395 400

Leu Glu Ser Ala Arg Ser Gly Asn Glu Glu Lys Met Met Ala Leu Leu
405 410 415

Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly Arg Lys Ser Thr
420 425 430

Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val Lys Ile Val Gln Leu
435 440 445

Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly Asp Leu
450 455 460

Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Thr Glu
465 470 475 480

Leu Leu Val Lys His Gly Ala Cys Val Asn Ala Met Asp Leu Trp Gln
485 490 495

Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val Glu Val Cys
500 505 510

Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr Leu Leu Asn Cys His
515 520 525

Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro Gln Leu Lys Glu Arg
530 535 540

Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala Ala Arg Glu
545 550 555 560

Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser Leu Glu Met Val Asn
565 570 575

Phe Lys His Pro Gln Thr His Glu Thr Ala Leu His Cys Ala Ala Ala
580 585 590

Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu Leu Leu Leu Arg Lys
595 600 605

Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe Leu Thr Pro Leu His
610 615 620

Val Ala Ser Glu Lys Ala His Asn Asp Val Val Glu Val Val Val Lys
625 630 635 640

His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu
645 650 655

His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu
660 665 670

Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala
675 680 685

Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile
690 695 700

Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys
705 710 715 720

Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val
725 730 735

Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala
740 745 750

Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu Gln His Gly
755 760 765

Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn
770 775 780

Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu Val Lys His
785 790 795 800

Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His
805 810 815

Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln
820 825 830

His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu
835 840 845

Asp Leu Val Lys Asp Gly Asp Thr Asp Ile Gln Asp Leu Leu Arg Gly
850 855 860

Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly Cys Leu Ala Arg Val
865 870 875 880

Lys Lys Leu Ser Ser Pro Asp Asn Val Asn Cys Arg Asp Thr Gln Gly
885 890 895

Arg His Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Leu Glu
900 905 910

Val Ala Glu Tyr Leu Leu Gln His Gly Ala Asp Val Asn Ala Gln Asp
915 920 925

Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala Ser Tyr Gly His Val
930 935 940

Asp Val Ala Ala Leu Leu Ile Lys Tyr Asn Ala Cys Val Asn Ala Thr
945 950 955 960

Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg
965 970 975

Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly Ala Asp Pro Thr Leu
980 985 990

Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu Val Ser Ala Asp Asp
995 1000 1005

Val Ser Ala Leu Leu Thr Ala Ala Met Pro Pro Ser Ala Leu Pro Ser
1010 1015 1020

Cys Tyr Lys Pro Gln Val Leu Asn Gly Val Arg Ser Pro Gly Ala Thr
1025 1030 1035 1040

Ala Asp Ala Leu Ser Ser Gly Pro Ser Ser Pro Ser Ser Leu Ser Ala
1045 1050 1055

Ala Ser Ser Leu Asp Asn Leu Ser Gly Ser Phe Ser Glu Leu Ser Ser
1060 1065 1070

```
Val Val Ser Ser Ser Gly Thr Glu Gly Ala Ser Ser Leu Glu Lys Lys
        1075                1080                1085

Glu Val Pro Gly Val Asp Phe Ser Ile Thr Gln Phe Val Arg Asn Leu
    1090                1095                1100

Gly Leu Glu His Leu Met Asp Ile Phe Glu Arg Glu Gln Ile Thr Leu
1105                1110                1115                1120

Asp Val Leu Val Glu Met Gly His Lys Glu Leu Lys Glu Ile Gly Ile
                1125                1130                1135

Asn Ala Tyr Gly His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu
            1140                1145                1150

Ile Ser Gly Gln Gln Gly Leu Asn Pro Tyr Leu Thr Leu Asn Thr Ser
        1155                1160                1165

Gly Ser Gly Thr Ile Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe
    1170                1175                1180

Gln Ser Val Glu Glu Glu Met Gln Ser Thr Val Arg Glu His Arg Asp
1185                1190                1195                1200

Gly Gly His Ala Gly Gly Ile Phe Asn Arg Tyr Asn Ile Leu Lys Ile
                1205                1210                1215

Gln Lys Val Cys Asn Lys Lys Leu Trp Glu Arg Tyr Thr His Arg Arg
            1220                1225                1230

Lys Glu Val Ser Glu Glu Asn His Asn His Ala Asn Glu Arg Met Leu
        1235                1240                1245

Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly Phe Asp
    1250                1255                1260

Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe
1265                1270                1275                1280

Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly
                1285                1290                1295

Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile Cys His
            1300                1305                1310

Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln
        1315                1320                1325

Phe Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His Ser Val
    1330                1335                1340

Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr Val Ile
1345                1350                1355                1360

Tyr Arg Gly Glu Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile
                1365                1370                1375

Met Arg Pro Glu Gly Met Val Asp Gly
            1380                1385
```

<210> 134
<211> 4992
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (876)..(4373)

```
<400> 134
ggaaagagta attgatcaga gccatccctc caattggagt caacttccat gactgttcgg 60

atttccttta ttttgggggc agttcatcca aacttctatt aaacggcaac tagttcactt 120

ttgagaagtg gtttacaaga aacaacaaca acaacaacaa agcagttgcg gaggaaagaa 180

aagagacaaa gtaaaaaaaa cggaaaagaa atctcccagg agaaagggat gtggaagctg 240

aaaacacgga caatttccac agtaagactt ccaaaagaat gtgcaagatc cgagcaaaac 300

tttcaagggc tctttttcag tgtaatggta gtgagaaagt tcagcctgga aagcccaggg 360

cttaaaacaa caacaacaaa aaacacaata tgcaggatcg ttcggcttca gcagaaccca 420

ccgcaaagat ggcggtggga cgaagcccct tctcccgccg ccgaagcctc tcgcctcaca 480

tttcccacaa acccttcgcg ccgcctcgct agccgaaacc tgcccagccg gtgcccggcc 540

actgcgcacg cgcgggacga cgtcacgtgc gctcccgggg ctggacggag ctggcaggag 600

gggccttgcc agcttccgcc gccgcgtcgt ttcaggaccc ggacggcgga ttcgcgctgc 660

ctccgccgcc gcggggcagc cggggggcag ggagcccagc gaggggcgcg cgtgggcgcg 720

gccatgggac tgcgccggat ccggtgacag cagggagcca agcggcccgg gccctgagcg 780

cgtcttctcc ggggggcctc gccctcctgc tcgcggggcc ggggctcctg ctccggttgc 840

tggcgctgtt gctggctgtg gcggcggcca ggatc atg tcg ggt cgc cgc tgc    893
                                       Met Ser Gly Arg Arg Cys
                                       1               5

gcc ggc ggg gga gcg gcc tgc gcg agc gcc gcg gcc gag gcc gtg gag    941
Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala Ala Ala Glu Ala Val Glu
            10                  15                  20

ccg gcc gcc cga gag ctg ttc gag gcg tgc cgc aac ggg gac gtg gaa    989
Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys Arg Asn Gly Asp Val Glu
        25                  30                  35

cga gtc aag agg ctg gtg acg cct gag aag gtg aac agc cgc gac acg    1037
Arg Val Lys Arg Leu Val Thr Pro Glu Lys Val Asn Ser Arg Asp Thr
    40                  45                  50

gcg ggc agg aaa tcc acc ccg ctg cac ttc gcc gca ggt ttt ggg cgg    1085
Ala Gly Arg Lys Ser Thr Pro Leu His Phe Ala Ala Gly Phe Gly Arg
55                  60                  65                  70

aaa gac gta gtt gaa tat ttg ctt cag aat ggt gca aat gtc caa gca    1133
Lys Asp Val Val Glu Tyr Leu Leu Gln Asn Gly Ala Asn Val Gln Ala
                75                  80                  85

cgt gat gat ggg ggc ctt att cct ctt cat aat gca tgc tct ttt ggt    1181
Arg Asp Asp Gly Gly Leu Ile Pro Leu His Asn Ala Cys Ser Phe Gly
            90                  95                  100

cat gct gaa gta gtc aat ctc ctt ttg cga cat ggt gca gac ccc aat    1229
His Ala Glu Val Val Asn Leu Leu Leu Arg His Gly Ala Asp Pro Asn
        105                 110                 115

gct cga gat aat tgg aat tat act cct ctc cat gaa gct gca att aaa    1277
Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu His Glu Ala Ala Ile Lys
    120                 125                 130

gga aag att gat gtt tgc att gtg ctg tta cag cat gga gct gag cca    1325
Gly Lys Ile Asp Val Cys Ile Val Leu Leu Gln His Gly Ala Glu Pro
135                 140                 145                 150
```

```
acc atc cga aat aca gat gga agg aca gca ttg gat tta gca gat cca     1373
Thr Ile Arg Asn Thr Asp Gly Arg Thr Ala Leu Asp Leu Ala Asp Pro
                155                 160                 165

tct gcc aaa gca gtg ctt act ggt gaa tat aag aaa gat gaa ctc tta     1421
Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr Lys Lys Asp Glu Leu Leu
            170                 175                 180

gaa agt gcc agg agt ggc aat gaa gaa aaa atg atg gct cta ctc aca     1469
Glu Ser Ala Arg Ser Gly Asn Glu Glu Lys Met Met Ala Leu Leu Thr
        185                 190                 195

cca tta aat gtc aac tgc cac gca agt gat ggc aga aag tca act cca     1517
Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly Arg Lys Ser Thr Pro
    200                 205                 210

tta cat ttg gca gca gga tat aac aga gta aag att gta cag ctg tta     1565
Leu His Leu Ala Ala Gly Tyr Asn Arg Val Lys Ile Val Gln Leu Leu
215                 220                 225                 230

ctg caa cat gga gct gat gtc cat gct aaa gat aaa ggt gat ctg gta     1613
Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly Asp Leu Val
                235                 240                 245

cca tta cac aat gcc tgt tct tat ggt cat tat gaa gta act gaa ctt     1661
Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Thr Glu Leu
            250                 255                 260

ttg gtc aag cat ggt gcc tgt gta aat gca atg gac ttg tgg caa ttc     1709
Leu Val Lys His Gly Ala Cys Val Asn Ala Met Asp Leu Trp Gln Phe
        265                 270                 275

act cct ctt cat gag gca gct tct aag aac agg gtt gaa gta tgt tct     1757
Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val Glu Val Cys Ser
    280                 285                 290

ctt ctc tta agt tat ggt gca gac cca aca ctg ctc aat tgt cac aat     1805
Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr Leu Leu Asn Cys His Asn
295                 300                 305                 310

aaa agt gct ata gac ttg gct ccc aca cca cag tta aaa gaa aga tta     1853
Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro Gln Leu Lys Glu Arg Leu
                315                 320                 325

gca tat gaa ttt aaa ggc cac tcg ttg ctg caa gct gca cga gaa gct     1901
Ala Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala Ala Arg Glu Ala
            330                 335                 340

gat gtt act cga atc aaa aaa cat ctc tct ctg gaa atg gtg aat ttc     1949
Asp Val Thr Arg Ile Lys Lys His Leu Ser Leu Glu Met Val Asn Phe
        345                 350                 355

aag cat cct caa aca cat gaa aca gca ttg cat tgt gct gct gca tct     1997
Lys His Pro Gln Thr His Glu Thr Ala Leu His Cys Ala Ala Ala Ser
    360                 365                 370

cca tat ccc aaa aga aag caa ata tgt gaa ctg ttg cta aga aaa gga     2045
Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu Leu Leu Leu Arg Lys Gly
375                 380                 385                 390

gca aac atc aat gaa aag act aaa gaa ttc ttg act cct ctg cac gtg     2093
Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe Leu Thr Pro Leu His Val
                395                 400                 405

gca tct gag aaa gct cat aat gat gtt gtt gaa gta gtg gtg aaa cat     2141
Ala Ser Glu Lys Ala His Asn Asp Val Val Glu Val Val Val Lys His
            410                 415                 420
```

```
gaa gca aag gtt aat gct ctg gat aat ctt ggt cag act tct cta cac     2189
Glu Ala Lys Val Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu His
        425                 430                 435

aga gct gca tat tgt ggt cat cta caa acc tgc cgc cta ctc ctg agc     2237
Arg Ala Ala Tyr Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu Ser
    440                 445                 450

tat ggg tgt gat cct aac att ata tcc ctt cag ggc ttt act gct tta     2285
Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala Leu
455                 460                 465                 470

cag atg gga aat gaa aat gta cag caa ctc ctc caa gag ggt atc tca     2333
Gln Met Gly Asn Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile Ser
                475                 480                 485

tta ggt aat tca gag gca gac aga caa ttg ctg gaa gct gca aag gct     2381
Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys Ala
            490                 495                 500

gga gat gtc gaa act gta aaa aaa ctg tgt act gtt cag agt gtc aac     2429
Gly Asp Val Glu Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val Asn
        505                 510                 515

tgc aga gac att gaa ggg cgt cag tct aca cca ctt cat ttt gca gct     2477
Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala Ala
    520                 525                 530

ggg tat aac aga gtg tcc gtg gtg gaa tat ctg cta cag cat gga gct     2525
Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu Gln His Gly Ala
535                 540                 545                 550

gat gtg cat gct aaa gat aaa gga ggc ctt gta cct ttg cac aat gca     2573
Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala
                555                 560                 565

tgt tct tat gga cat tat gaa gtt gca gaa ctt ctt gtt aaa cat gga     2621
Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu Val Lys His Gly
            570                 575                 580

gca gta gtt aat gta gct gat tta tgg aaa ttt aca cct tta cat gaa     2669
Ala Val Val Asn Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His Glu
        585                 590                 595

gca gca gca aaa gga aaa tat gaa att tgc aaa ctt ctg ctc cag cat     2717
Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln His
    600                 605                 610

ggt gca gac cct aca aaa aaa aac agg gat gga aat act cct ttg gat     2765
Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu Asp
615                 620                 625                 630

ctt gtt aaa gat gga gat aca gat att caa gat ctg ctt agg gga gat     2813
Leu Val Lys Asp Gly Asp Thr Asp Ile Gln Asp Leu Leu Arg Gly Asp
                635                 640                 645

gca gct ttg cta gat gct gcc aag aag ggt tgt tta gcc aga gtg aag     2861
Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly Cys Leu Ala Arg Val Lys
            650                 655                 660

aag ttg tct tct cct gat aat gta aat tgc cgc gat acc caa ggc aga     2909
Lys Leu Ser Ser Pro Asp Asn Val Asn Cys Arg Asp Thr Gln Gly Arg
        665                 670                 675

cat tca aca cct tta cat tta gca gct ggt tat aat aat tta gaa gtt     2957
His Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Leu Glu Val
    680                 685                 690
```

```
gca gag tat ttg tta caa cac gga gct gat gtg aat gcc caa gac aaa     3005
Ala Glu Tyr Leu Leu Gln His Gly Ala Asp Val Asn Ala Gln Asp Lys
695                 700                 705                 710

gga gga ctt att cct tta cat aat gca gca tct tac ggg cat gta gat     3053
Gly Gly Leu Ile Pro Leu His Asn Ala Ala Ser Tyr Gly His Val Asp
                715                 720                 725

gta gca gct cta cta ata aag tat aat gca tgt gtc aat gcc acg gac     3101
Val Ala Ala Leu Leu Ile Lys Tyr Asn Ala Cys Val Asn Ala Thr Asp
            730                 735                 740

aaa tgg gct ttc aca cct ttg cac gaa gca gcc caa aag gga cga aca     3149
Lys Trp Ala Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr
        745                 750                 755

cag ctt tgt gct ttg ttg cta gcc cat gga gct gac ccg act ctt aaa     3197
Gln Leu Cys Ala Leu Leu Leu Ala His Gly Ala Asp Pro Thr Leu Lys
    760                 765                 770

aat cag gaa gga caa aca cct tta gat tta gtt tca gca gat gat gtc     3245
Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu Val Ser Ala Asp Asp Val
775                 780                 785                 790

agc gct ctt ctg aca gca gcc atg ccc cca tct gct ctg ccc tct tgt     3293
Ser Ala Leu Leu Thr Ala Ala Met Pro Pro Ser Ala Leu Pro Ser Cys
                795                 800                 805

tac aag cct caa gtg ctc aat ggt gtg aga agc cca gga gcc act gca     3341
Tyr Lys Pro Gln Val Leu Asn Gly Val Arg Ser Pro Gly Ala Thr Ala
            810                 815                 820

gat gct ctc tct tca ggt cca tct agc cca tca agc ctt tct gca gcc     3389
Asp Ala Leu Ser Ser Gly Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala
        825                 830                 835

agc agt ctt gac aac tta tct ggg agt ttt tca gaa ctg tct tca gta     3437
Ser Ser Leu Asp Asn Leu Ser Gly Ser Phe Ser Glu Leu Ser Ser Val
    840                 845                 850

gtt agt tca agt gga aca gag ggt gct tcc agt ttg gag aaa aag gag     3485
Val Ser Ser Ser Gly Thr Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu
855                 860                 865                 870

gtt cca gga gta gat ttt agc ata act caa ttc gta agg aat ctt gga     3533
Val Pro Gly Val Asp Phe Ser Ile Thr Gln Phe Val Arg Asn Leu Gly
                875                 880                 885

ctt gag cac cta atg gat ata ttt gag aga gaa cag atc act ttg gat     3581
Leu Glu His Leu Met Asp Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp
            890                 895                 900

gta tta gtt gag atg ggg cac aag gag ctg aag gag att gga atc aat     3629
Val Leu Val Glu Met Gly His Lys Glu Leu Lys Glu Ile Gly Ile Asn
        905                 910                 915

gct tat gga cat agg cac aaa cta att aaa gga gtc gag aga ctt atc     3677
Ala Tyr Gly His Arg His Lys Leu Ile Lys Gly Val Glu Arg Leu Ile
    920                 925                 930

tcc gga caa caa ggt ctt aac cca tat tta act ttg aac acc tct ggt     3725
Ser Gly Gln Gln Gly Leu Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly
935                 940                 945                 950

agt gga aca att ctt ata gat ctg tct cct gat gat aaa gag ttt cag     3773
Ser Gly Thr Ile Leu Ile Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln
                955                 960                 965
```

tct gtg gag gaa gag atg caa agt aca gtt cga gag cac aga gat gga 3821
Ser Val Glu Glu Glu Met Gln Ser Thr Val Arg Glu His Arg Asp Gly
970 975 980

ggt cat gca ggt gga atc ttc aac aga tac aat att ctc aag att cag 3869
Gly His Ala Gly Gly Ile Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln
985 990 995

aag gtt tgt aac aag aaa cta tgg gaa aga tac act cac cgg aga aaa 3917
Lys Val Cys Asn Lys Lys Leu Trp Glu Arg Tyr Thr His Arg Arg Lys
1000 1005 1010

gaa gtt tct gaa gaa aac cac aac cat gcc aat gaa cga atg cta ttt 3965
Glu Val Ser Glu Glu Asn His Asn His Ala Asn Glu Arg Met Leu Phe
1015 1020 1025 1030

cat ggg tct cct ttt gtg aat gca att atc cac aaa ggc ttt gat gaa 4013
His Gly Ser Pro Phe Val Asn Ala Ile Ile His Lys Gly Phe Asp Glu
1035 1040 1045

agg cat gcg tac ata ggt ggt atg ttt gga gct ggc att tat ttt gct 4061
Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala
1050 1055 1060

gaa aac tct tcc aaa agc aat caa tat gta tat gga att gga gga ggt 4109
Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly
1065 1070 1075

act ggg tgt cca gtt cac aaa gac aga tct tgt tac att tgc cac agg 4157
Thr Gly Cys Pro Val His Lys Asp Arg Ser Cys Tyr Ile Cys His Arg
1080 1085 1090

cag ctg ctc ttt tgc cgg gta acc ttg gga aag tct ttc ctg cag ttc 4205
Gln Leu Leu Phe Cys Arg Val Thr Leu Gly Lys Ser Phe Leu Gln Phe
1095 1100 1105 1110

agt gca atg aaa atg gca cat tct cct cca ggt cat cac tca gtc act 4253
Ser Ala Met Lys Met Ala His Ser Pro Pro Gly His His Ser Val Thr
1115 1120 1125

ggt agg ccc agt gta aat ggc cta gca tta gct gaa tat gtt att tac 4301
Gly Arg Pro Ser Val Asn Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr
1130 1135 1140

aga gga gaa cag gct tat cct gag tat tta att act tac cag att atg 4349
Arg Gly Glu Gln Ala Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met
1145 1150 1155

agg cct gaa ggt atg gtc gat gga taaatagtta ttttaagaaa ctaattccac 4403
Arg Pro Glu Gly Met Val Asp Gly
1160 1165

tgaacctaaa atcatcaaag cagcagtggc ctctacgttt tactcctttg ctgaaaaaaa 4463

atcatcttgc ccacaggcct gtggcaaaag gataaaaatg tgaacgaagt ttaacattct 4523

gacttgataa agctttaata atgtacagtg ttttctaaat atttcctgtt ttttcagcac 4583

tttaacagat gccattccag gttaaactgg gttgtctgta ctaaattata aacagagtta 4643

acttgaacct tttatatgtt atgcattgat tctaacaaac tgtaatgccc tcaacagaac 4703

taattttact aatacaatac tgtgttcttt aaaacacagc atttacactg aatacaattt 4763

catttgtaaa actgtaaata agagcttttg tactagccca gtatttattt acattgcttt 4823

gtaatataaa tctgttttag aactgcagcg gtttacaaaa ttttttcata tgtattgttc 4883

atctatactt catcttacat cgtcatgatt gagtgatctt tacatttgat tccagaggct 4943

atgttcagtt gttagttggg aaagattgag ttatcagatt taatttgcc 4992

<210> 135
<211> 1166
<212> PRT
<213> Homo sapiens

<400> 135
Met Ser Gly Arg Arg Cys Ala Gly Gly Gly Ala Ala Cys Ala Ser Ala
1 5 10 15

Ala Ala Glu Ala Val Glu Pro Ala Ala Arg Glu Leu Phe Glu Ala Cys
20 25 30

Arg Asn Gly Asp Val Glu Arg Val Lys Arg Leu Val Thr Pro Glu Lys
35 40 45

Val Asn Ser Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe
50 55 60

Ala Ala Gly Phe Gly Arg Lys Asp Val Val Glu Tyr Leu Leu Gln Asn
65 70 75 80

Gly Ala Asn Val Gln Ala Arg Asp Asp Gly Gly Leu Ile Pro Leu His
85 90 95

Asn Ala Cys Ser Phe Gly His Ala Glu Val Val Asn Leu Leu Leu Arg
100 105 110

His Gly Ala Asp Pro Asn Ala Arg Asp Asn Trp Asn Tyr Thr Pro Leu
115 120 125

His Glu Ala Ala Ile Lys Gly Lys Ile Asp Val Cys Ile Val Leu Leu
130 135 140

Gln His Gly Ala Glu Pro Thr Ile Arg Asn Thr Asp Gly Arg Thr Ala
145 150 155 160

Leu Asp Leu Ala Asp Pro Ser Ala Lys Ala Val Leu Thr Gly Glu Tyr
165 170 175

Lys Lys Asp Glu Leu Leu Glu Ser Ala Arg Ser Gly Asn Glu Glu Lys
180 185 190

Met Met Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp
195 200 205

Gly Arg Lys Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Val
210 215 220

Lys Ile Val Gln Leu Leu Leu Gln His Gly Ala Asp Val His Ala Lys
225 230 235 240

Asp Lys Gly Asp Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His
245 250 255

Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala Cys Val Asn Ala
260 265 270

Met Asp Leu Trp Gln Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn
275 280 285

Arg Val Glu Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr
290 295 300

Leu Leu Asn Cys His Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro
305 310 315 320

Gln Leu Lys Glu Arg Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu
325 330 335

Gln Ala Ala Arg Glu Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser
340 345 350

Leu Glu Met Val Asn Phe Lys His Pro Gln Thr His Glu Thr Ala Leu
355 360 365

His Cys Ala Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu
370 375 380

Leu Leu Leu Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe
385 390 395 400

Leu Thr Pro Leu His Val Ala Ser Glu Lys Ala His Asn Asp Val Val
405 410 415

Glu Val Val Val Lys His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu
420 425 430

Gly Gln Thr Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr
435 440 445

Cys Arg Leu Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu
450 455 460

Gln Gly Phe Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu
465 470 475 480

Leu Gln Glu Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu
485 490 495

Leu Glu Ala Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys
500 505 510

Thr Val Gln Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr
515 520 525

Pro Leu His Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr
530 535 540

Leu Leu Gln His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu
545 550 555 560

Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu
565 570 575

Leu Leu Val Lys His Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys
580 585 590

Phe Thr Pro Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys
595 600 605

Lys Leu Leu Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp
610 615 620

Gly Asn Thr Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp Ile Gln
625 630 635 640

Asp Leu Leu Arg Gly Asp Ala Ala Leu Leu Asp Ala Ala Lys Lys Gly
645 650 655

Cys Leu Ala Arg Val Lys Lys Leu Ser Ser Pro Asp Asn Val Asn Cys
660 665 670

Arg Asp Thr Gln Gly Arg His Ser Thr Pro Leu His Leu Ala Ala Gly
675 680 685

Tyr Asn Asn Leu Glu Val Ala Glu Tyr Leu Leu Gln His Gly Ala Asp
690 695 700

Val Asn Ala Gln Asp Lys Gly Gly Leu Ile Pro Leu His Asn Ala Ala
705 710 715 720

Ser Tyr Gly His Val Asp Val Ala Ala Leu Leu Ile Lys Tyr Asn Ala
725 730 735

Cys Val Asn Ala Thr Asp Lys Trp Ala Phe Thr Pro Leu His Glu Ala
740 745 750

Ala Gln Lys Gly Arg Thr Gln Leu Cys Ala Leu Leu Leu Ala His Gly
755 760 765

Ala Asp Pro Thr Leu Lys Asn Gln Glu Gly Gln Thr Pro Leu Asp Leu
770 775 780

Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala Ala Met Pro Pro
785 790 795 800

Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu Asn Gly Val Arg
805 810 815

Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly Pro Ser Ser Pro
820 825 830

Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu Ser Gly Ser Phe
835 840 845

Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr Glu Gly Ala Ser
850 855 860

Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe Ser Ile Thr Gln
865 870 875 880

Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp Ile Phe Glu Arg
885 890 895

Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly His Lys Glu Leu
900 905 910

Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His Lys Leu Ile Lys
915 920 925

Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu Asn Pro Tyr Leu
930 935 940

Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile Asp Leu Ser Pro
945 950 955 960

Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met Gln Ser Thr Val
965 970 975

Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile Phe Asn Arg Tyr
980 985 990

Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys Leu Trp Glu Arg
995 1000 1005

Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn His Asn His Ala
1010 1015 1020

Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val Asn Ala Ile Ile
1025 1030 1035 1040

His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly Gly Met Phe Gly
1045 1050 1055

```
Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser Asn Gln Tyr Val
        1060                1065                1070

Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His Lys Asp Arg Ser
    1075                1080                1085

Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg Val Thr Leu Gly
    1090                1095                1100

Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala His Ser Pro Pro
1105                1110                1115                1120

Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn Gly Leu Ala Leu
                1125                1130                1135

Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr Pro Glu Tyr Leu
        1140                1145                1150

Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val Asp Gly
    1155                1160                1165


<210> 136
<211> 3045
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(3042)

<400> 136
atg gcg gag tct tcg gat aag ctc tat cga gtc gag tac gcc aag agc      48
Met Ala Glu Ser Ser Asp Lys Leu Tyr Arg Val Glu Tyr Ala Lys Ser
1               5                   10                  15

ggg cgc gcc tct tgc aag aaa tgc agc gag agc atc ccc aag gac tcg      96
Gly Arg Ala Ser Cys Lys Lys Cys Ser Glu Ser Ile Pro Lys Asp Ser
            20                  25                  30

ctc cgg atg gcc atc atg gtg cag tcg ccc atg ttt gat gga aaa gtc     144
Leu Arg Met Ala Ile Met Val Gln Ser Pro Met Phe Asp Gly Lys Val
        35                  40                  45

cca cac tgg tac cac ttc tcc tgc ttc tgg aag gtg ggc cac tcc atc     192
Pro His Trp Tyr His Phe Ser Cys Phe Trp Lys Val Gly His Ser Ile
    50                  55                  60

cgg cac cct gac gtt gag gtg gat ggg ttc tct gag ctt cgg tgg gat     240
Arg His Pro Asp Val Glu Val Asp Gly Phe Ser Glu Leu Arg Trp Asp
65                  70                  75                  80

gac cag cag aaa gtc aag aag aca gcg gaa gct gga gga gtg aca ggc     288
Asp Gln Gln Lys Val Lys Lys Thr Ala Glu Ala Gly Gly Val Thr Gly
                85                  90                  95

aaa ggc cag gat gga att ggt agc aag gca gag aag act ctg ggt gac     336
Lys Gly Gln Asp Gly Ile Gly Ser Lys Ala Glu Lys Thr Leu Gly Asp
            100                 105                 110

ttt gca gca gag tat gcc aag tcc aac aga agt acg tgc aag ggg tgt     384
Phe Ala Ala Glu Tyr Ala Lys Ser Asn Arg Ser Thr Cys Lys Gly Cys
        115                 120                 125

atg gag aag ata gaa aag ggc cag gtg cgc ctg tcc aag aag atg gtg     432
Met Glu Lys Ile Glu Lys Gly Gln Val Arg Leu Ser Lys Lys Met Val
    130                 135                 140
```

gac ccg gag aag cca cag cta ggc atg att gac cgc tgg tac cat cca 480
Asp Pro Glu Lys Pro Gln Leu Gly Met Ile Asp Arg Trp Tyr His Pro
145 150 155 160

ggc tgc ttt gtc aag aac agg gag gag ctg ggt ttc cgg ccc gag tac 528
Gly Cys Phe Val Lys Asn Arg Glu Glu Leu Gly Phe Arg Pro Glu Tyr
165 170 175

agt gcg agt cag ctc aag ggc ttc agc ctc ctt gct aca gag gat aaa 576
Ser Ala Ser Gln Leu Lys Gly Phe Ser Leu Leu Ala Thr Glu Asp Lys
180 185 190

gaa gcc ctg aag aag cag ctc cca gga gtc aag agt gaa gga aag aga 624
Glu Ala Leu Lys Lys Gln Leu Pro Gly Val Lys Ser Glu Gly Lys Arg
195 200 205

aaa ggc gat gag gtg gat gga gtg gat gaa gtg gcg aag aag aaa tct 672
Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys Lys Ser
210 215 220

aaa aaa gaa aaa gac aag gat agt aag ctt gaa aaa gcc cta aag gct 720
Lys Lys Glu Lys Asp Lys Asp Ser Lys Leu Glu Lys Ala Leu Lys Ala
225 230 235 240

cag aac gac ctg atc tgg aac atc aag gac gag cta aag aaa gtg tgt 768
Gln Asn Asp Leu Ile Trp Asn Ile Lys Asp Glu Leu Lys Lys Val Cys
245 250 255

tca act aat gac ctg aag gag cta ctc atc ttc aac aag cag caa gtg 816
Ser Thr Asn Asp Leu Lys Glu Leu Leu Ile Phe Asn Lys Gln Gln Val
260 265 270

cct tct ggg gag tcg gcg atc ttg gac cga gta gct gat ggc atg gtg 864
Pro Ser Gly Glu Ser Ala Ile Leu Asp Arg Val Ala Asp Gly Met Val
275 280 285

ttc ggt gcc ctc ctt ccc tgc gag gaa tgc tcg ggt cag ctg gtc ttc 912
Phe Gly Ala Leu Leu Pro Cys Glu Glu Cys Ser Gly Gln Leu Val Phe
290 295 300

aag agc gat gcc tat tac tgc act ggg gac gtc act gcc tgg acc aag 960
Lys Ser Asp Ala Tyr Tyr Cys Thr Gly Asp Val Thr Ala Trp Thr Lys
305 310 315 320

tgt atg gtc aag aca cag aca ccc aac cgg aag gag tgg gta acc cca 1008
Cys Met Val Lys Thr Gln Thr Pro Asn Arg Lys Glu Trp Val Thr Pro
325 330 335

aag gaa ttc cga gaa atc tct tac ctc aag aaa ttg aag gtt aaa aag 1056
Lys Glu Phe Arg Glu Ile Ser Tyr Leu Lys Lys Leu Lys Val Lys Lys
340 345 350

cag gac cgt ata ttc ccc cca gaa acc agc gcc tcc gtg gcg gcc acg 1104
Gln Asp Arg Ile Phe Pro Pro Glu Thr Ser Ala Ser Val Ala Ala Thr
355 360 365

cct ccg ccc tcc aca gcc tcg gct cct gct gct gtg aac tcc tct gct 1152
Pro Pro Pro Ser Thr Ala Ser Ala Pro Ala Ala Val Asn Ser Ser Ala
370 375 380

tca gca gat aag cca tta tcc aac atg aag atc ctg act ctc ggg aag 1200
Ser Ala Asp Lys Pro Leu Ser Asn Met Lys Ile Leu Thr Leu Gly Lys
385 390 395 400

ctg tcc cgg aac aag gat gaa gtg aag gcc atg att gag aaa ctc ggg 1248
Leu Ser Arg Asn Lys Asp Glu Val Lys Ala Met Ile Glu Lys Leu Gly
405 410 415

```
ggg aag ttg acg ggg acg gcc aac aag gct tcc ctg tgc atc agc acc     1296
Gly Lys Leu Thr Gly Thr Ala Asn Lys Ala Ser Leu Cys Ile Ser Thr
            420                 425                 430

aaa aag gag gtg gaa aag atg aat aag aag atg gag gaa gta aag gaa     1344
Lys Lys Glu Val Glu Lys Met Asn Lys Lys Met Glu Glu Val Lys Glu
        435                 440                 445

gcc aac atc cga gtt gtg tct gag gac ttc ctc cag gac gtc tcc gcc     1392
Ala Asn Ile Arg Val Val Ser Glu Asp Phe Leu Gln Asp Val Ser Ala
    450                 455                 460

tcc acc aag agc ctt cag gag ttg ttc tta gcg cac atc ttg tcc cct     1440
Ser Thr Lys Ser Leu Gln Glu Leu Phe Leu Ala His Ile Leu Ser Pro
465                 470                 475                 480

tgg ggg gca gag gtg aag gca gag cct gtt gaa gtt gtg gcc cca aga     1488
Trp Gly Ala Glu Val Lys Ala Glu Pro Val Glu Val Val Ala Pro Arg
                485                 490                 495

ggg aag tca ggg gct gcg ctc tcc aaa aaa agc aag ggc cag gtc aag     1536
Gly Lys Ser Gly Ala Ala Leu Ser Lys Lys Ser Lys Gly Gln Val Lys
            500                 505                 510

gag gaa ggt atc aac aaa tct gaa aag aga atg aaa tta act ctt aaa     1584
Glu Glu Gly Ile Asn Lys Ser Glu Lys Arg Met Lys Leu Thr Leu Lys
        515                 520                 525

gga gga gca gct gtg gat cct gat tct gga ctg gaa cac tct gcg cat     1632
Gly Gly Ala Ala Val Asp Pro Asp Ser Gly Leu Glu His Ser Ala His
    530                 535                 540

gtc ctg gag aaa ggt ggg aag gtc ttc agt gcc acc ctt ggc ctg gtg     1680
Val Leu Glu Lys Gly Gly Lys Val Phe Ser Ala Thr Leu Gly Leu Val
545                 550                 555                 560

gac atc gtt aaa gga acc aac tcc tac tac aag ctg cag ctt ctg gag     1728
Asp Ile Val Lys Gly Thr Asn Ser Tyr Tyr Lys Leu Gln Leu Leu Glu
                565                 570                 575

gac gac aag gaa aac agg tat tgg ata ttc agg tcc tgg ggc cgt gtg     1776
Asp Asp Lys Glu Asn Arg Tyr Trp Ile Phe Arg Ser Trp Gly Arg Val
            580                 585                 590

ggt acg gtg atc ggt agc aac aaa ctg gaa cag atg ccg tcc aag gag     1824
Gly Thr Val Ile Gly Ser Asn Lys Leu Glu Gln Met Pro Ser Lys Glu
        595                 600                 605

gat gcc att gag cag ttc atg aaa tta tat gaa gaa aaa acc ggg aac     1872
Asp Ala Ile Glu Gln Phe Met Lys Leu Tyr Glu Glu Lys Thr Gly Asn
    610                 615                 620

gct tgg cac tcc aaa aat ttc acg aag tat ccc aaa aag ttt tac ccc     1920
Ala Trp His Ser Lys Asn Phe Thr Lys Tyr Pro Lys Lys Phe Tyr Pro
625                 630                 635                 640

ctg gag att gac tat ggc cag gat gaa gag gca gtg aag aag ctc aca     1968
Leu Glu Ile Asp Tyr Gly Gln Asp Glu Glu Ala Val Lys Lys Leu Thr
                645                 650                 655

gta aat cct ggc acc aag tcc aag ctc ccc aag cca gtt cag gac ctc     2016
Val Asn Pro Gly Thr Lys Ser Lys Leu Pro Lys Pro Val Gln Asp Leu
            660                 665                 670

atc aag atg atc ttt gat gtg gaa agt atg aag aaa gcc atg gtg gag     2064
Ile Lys Met Ile Phe Asp Val Glu Ser Met Lys Lys Ala Met Val Glu
        675                 680                 685
```

```
tat gag atc gac ctt cag aag atg ccc ttg ggg aag ctg agc aaa agg     2112
Tyr Glu Ile Asp Leu Gln Lys Met Pro Leu Gly Lys Leu Ser Lys Arg
    690                 695                 700

cag atc cag gcc gca tac tcc atc ctc agt gag gtc cag cag gcg gtg     2160
Gln Ile Gln Ala Ala Tyr Ser Ile Leu Ser Glu Val Gln Gln Ala Val
705                 710                 715                 720

tct cag ggc agc agc gac tct cag atc ctg gat ctc tca aat cgc ttt     2208
Ser Gln Gly Ser Ser Asp Ser Gln Ile Leu Asp Leu Ser Asn Arg Phe
                725                 730                 735

tac acc ctg atc ccc cac gac ttt ggg atg aag aag cct ccg ctc ctg     2256
Tyr Thr Leu Ile Pro His Asp Phe Gly Met Lys Lys Pro Pro Leu Leu
            740                 745                 750

aac aat gca gac agt gtg cag gcc aag gtg gaa atg ctt gac aac ctg     2304
Asn Asn Ala Asp Ser Val Gln Ala Lys Val Glu Met Leu Asp Asn Leu
        755                 760                 765

ctg gac atc gag gtg gcc tac agt ctg ctc agg gga ggg tct gat gat     2352
Leu Asp Ile Glu Val Ala Tyr Ser Leu Leu Arg Gly Gly Ser Asp Asp
    770                 775                 780

agc agc aag gat ccc atc gat gtc aac tat gag aag ctc aaa act gac     2400
Ser Ser Lys Asp Pro Ile Asp Val Asn Tyr Glu Lys Leu Lys Thr Asp
785                 790                 795                 800

att aag gtg gtt gac aga gat tct gaa gaa gcc gag atc atc agg aag     2448
Ile Lys Val Val Asp Arg Asp Ser Glu Glu Ala Glu Ile Ile Arg Lys
                805                 810                 815

tat gtt aag aac act cat gca acc aca cac agt gcg tat gac ttg gaa     2496
Tyr Val Lys Asn Thr His Ala Thr Thr His Ser Ala Tyr Asp Leu Glu
            820                 825                 830

gtc atc gat atc ttt aag ata gag cgt gaa ggc gaa tgc cag cgt tac     2544
Val Ile Asp Ile Phe Lys Ile Glu Arg Glu Gly Glu Cys Gln Arg Tyr
        835                 840                 845

aag ccc ttt aag cag ctt cat aac cga aga ttg ctg tgg cac ggg tcc     2592
Lys Pro Phe Lys Gln Leu His Asn Arg Arg Leu Leu Trp His Gly Ser
    850                 855                 860

agg acc acc aac ttt gct ggg atc ctg tcc cag ggt ctt cgg ata gcc     2640
Arg Thr Thr Asn Phe Ala Gly Ile Leu Ser Gln Gly Leu Arg Ile Ala
865                 870                 875                 880

ccg cct gaa gcg ccc gtg aca ggc tac atg ttt ggt aaa ggg atc tat     2688
Pro Pro Glu Ala Pro Val Thr Gly Tyr Met Phe Gly Lys Gly Ile Tyr
                885                 890                 895

ttc gct gac atg gtc tcc aag agt gcc aac tac tac cat acg tct cag     2736
Phe Ala Asp Met Val Ser Lys Ser Ala Asn Tyr Tyr His Thr Ser Gln
            900                 905                 910

gga gac cca ata ggc tta atc ctg ttg gga gaa gtt gcc ctt gga aac     2784
Gly Asp Pro Ile Gly Leu Ile Leu Leu Gly Glu Val Ala Leu Gly Asn
        915                 920                 925

atg tat gaa ctg aag cac gct tca cat atc agc agg tta ccc aag ggc     2832
Met Tyr Glu Leu Lys His Ala Ser His Ile Ser Arg Leu Pro Lys Gly
    930                 935                 940

aag cac agt gtc aaa ggt ttg ggc aaa act acc cct gat cct tca gct     2880
Lys His Ser Val Lys Gly Leu Gly Lys Thr Thr Pro Asp Pro Ser Ala
945                 950                 955                 960
```

```
aac att agt ctg gat ggt gta gac gtt cct ctt ggg acc ggg att tca     2928
Asn Ile Ser Leu Asp Gly Val Asp Val Pro Leu Gly Thr Gly Ile Ser
                965                 970                 975

tct ggt gtg ata gac acc tct cta cta tat aac gag tac att gtc tat     2976
Ser Gly Val Ile Asp Thr Ser Leu Leu Tyr Asn Glu Tyr Ile Val Tyr
            980                 985                 990

gat att gct cag gta aat ctg aag tat ctg ctg aaa ctg aaa ttc aat     3024
Asp Ile Ala Gln Val Asn Leu Lys Tyr Leu Leu Lys Leu Lys Phe Asn
        995                 1000                1005

ttt aag acc tcc ctg tgg taa                                         3045
Phe Lys Thr Ser Leu Trp
    1010


<210> 137
<211> 1014
<212> PRT
<213> Homo sapiens

<400> 137
Met Ala Glu Ser Ser Asp Lys Leu Tyr Arg Val Glu Tyr Ala Lys Ser
1               5                   10                  15

Gly Arg Ala Ser Cys Lys Lys Cys Ser Glu Ser Ile Pro Lys Asp Ser
            20                  25                  30

Leu Arg Met Ala Ile Met Val Gln Ser Pro Met Phe Asp Gly Lys Val
        35                  40                  45

Pro His Trp Tyr His Phe Ser Cys Phe Trp Lys Val Gly His Ser Ile
    50                  55                  60

Arg His Pro Asp Val Glu Val Asp Gly Phe Ser Glu Leu Arg Trp Asp
65                  70                  75                  80

Asp Gln Gln Lys Val Lys Lys Thr Ala Glu Ala Gly Gly Val Thr Gly
                85                  90                  95

Lys Gly Gln Asp Gly Ile Gly Ser Lys Ala Glu Lys Thr Leu Gly Asp
            100                 105                 110

Phe Ala Ala Glu Tyr Ala Lys Ser Asn Arg Ser Thr Cys Lys Gly Cys
        115                 120                 125

Met Glu Lys Ile Glu Lys Gly Gln Val Arg Leu Ser Lys Lys Met Val
    130                 135                 140

Asp Pro Glu Lys Pro Gln Leu Gly Met Ile Asp Arg Trp Tyr His Pro
145                 150                 155                 160

Gly Cys Phe Val Lys Asn Arg Glu Glu Leu Gly Phe Arg Pro Glu Tyr
                165                 170                 175

Ser Ala Ser Gln Leu Lys Gly Phe Ser Leu Leu Ala Thr Glu Asp Lys
            180                 185                 190

Glu Ala Leu Lys Lys Gln Leu Pro Gly Val Lys Ser Glu Gly Lys Arg
        195                 200                 205

Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys Lys Ser
    210                 215                 220

Lys Lys Glu Lys Asp Lys Asp Ser Lys Leu Glu Lys Ala Leu Lys Ala
225                 230                 235                 240

Gln Asn Asp Leu Ile Trp Asn Ile Lys Asp Glu Leu Lys Lys Val Cys
                245                 250                 255
```

```
Ser Thr Asn Asp Leu Lys Glu Leu Leu Ile Phe Asn Lys Gln Gln Val
            260                 265                 270

Pro Ser Gly Glu Ser Ala Ile Leu Asp Arg Val Ala Asp Gly Met Val
        275                 280                 285

Phe Gly Ala Leu Leu Pro Cys Glu Glu Cys Ser Gly Gln Leu Val Phe
    290                 295                 300

Lys Ser Asp Ala Tyr Tyr Cys Thr Gly Asp Val Thr Ala Trp Thr Lys
305                 310                 315                 320

Cys Met Val Lys Thr Gln Thr Pro Asn Arg Lys Glu Trp Val Thr Pro
                325                 330                 335

Lys Glu Phe Arg Glu Ile Ser Tyr Leu Lys Lys Leu Lys Val Lys Lys
            340                 345                 350

Gln Asp Arg Ile Phe Pro Pro Glu Thr Ser Ala Ser Val Ala Ala Thr
        355                 360                 365

Pro Pro Pro Ser Thr Ala Ser Ala Pro Ala Ala Val Asn Ser Ser Ala
    370                 375                 380

Ser Ala Asp Lys Pro Leu Ser Asn Met Lys Ile Leu Thr Leu Gly Lys
385                 390                 395                 400

Leu Ser Arg Asn Lys Asp Glu Val Lys Ala Met Ile Glu Lys Leu Gly
                405                 410                 415

Gly Lys Leu Thr Gly Thr Ala Asn Lys Ala Ser Leu Cys Ile Ser Thr
            420                 425                 430

Lys Lys Glu Val Glu Lys Met Asn Lys Lys Met Glu Glu Val Lys Glu
        435                 440                 445

Ala Asn Ile Arg Val Val Ser Glu Asp Phe Leu Gln Asp Val Ser Ala
    450                 455                 460

Ser Thr Lys Ser Leu Gln Glu Leu Phe Leu Ala His Ile Leu Ser Pro
465                 470                 475                 480

Trp Gly Ala Glu Val Lys Ala Glu Pro Val Glu Val Val Ala Pro Arg
                485                 490                 495

Gly Lys Ser Gly Ala Ala Leu Ser Lys Lys Ser Lys Gly Gln Val Lys
            500                 505                 510

Glu Glu Gly Ile Asn Lys Ser Glu Lys Arg Met Lys Leu Thr Leu Lys
        515                 520                 525

Gly Gly Ala Ala Val Asp Pro Asp Ser Gly Leu Glu His Ser Ala His
    530                 535                 540

Val Leu Glu Lys Gly Gly Lys Val Phe Ser Ala Thr Leu Gly Leu Val
545                 550                 555                 560

Asp Ile Val Lys Gly Thr Asn Ser Tyr Tyr Lys Leu Gln Leu Leu Glu
                565                 570                 575

Asp Asp Lys Glu Asn Arg Tyr Trp Ile Phe Arg Ser Trp Gly Arg Val
            580                 585                 590

Gly Thr Val Ile Gly Ser Asn Lys Leu Glu Gln Met Pro Ser Lys Glu
        595                 600                 605

Asp Ala Ile Glu Gln Phe Met Lys Leu Tyr Glu Glu Lys Thr Gly Asn
    610                 615                 620
```

```
Ala Trp His Ser Lys Asn Phe Thr Lys Tyr Pro Lys Lys Phe Tyr Pro
625                 630                 635                 640

Leu Glu Ile Asp Tyr Gly Gln Asp Glu Glu Ala Val Lys Lys Leu Thr
                645                 650                 655

Val Asn Pro Gly Thr Lys Ser Lys Leu Pro Lys Pro Val Gln Asp Leu
            660                 665                 670

Ile Lys Met Ile Phe Asp Val Glu Ser Met Lys Lys Ala Met Val Glu
        675                 680                 685

Tyr Glu Ile Asp Leu Gln Lys Met Pro Leu Gly Lys Leu Ser Lys Arg
    690                 695                 700

Gln Ile Gln Ala Ala Tyr Ser Ile Leu Ser Glu Val Gln Gln Ala Val
705                 710                 715                 720

Ser Gln Gly Ser Ser Asp Ser Gln Ile Leu Asp Leu Ser Asn Arg Phe
                725                 730                 735

Tyr Thr Leu Ile Pro His Asp Phe Gly Met Lys Lys Pro Pro Leu Leu
            740                 745                 750

Asn Asn Ala Asp Ser Val Gln Ala Lys Val Glu Met Leu Asp Asn Leu
        755                 760                 765

Leu Asp Ile Glu Val Ala Tyr Ser Leu Leu Arg Gly Gly Ser Asp Asp
    770                 775                 780

Ser Ser Lys Asp Pro Ile Asp Val Asn Tyr Glu Lys Leu Lys Thr Asp
785                 790                 795                 800

Ile Lys Val Val Asp Arg Asp Ser Glu Glu Ala Glu Ile Ile Arg Lys
                805                 810                 815

Tyr Val Lys Asn Thr His Ala Thr Thr His Ser Ala Tyr Asp Leu Glu
            820                 825                 830

Val Ile Asp Ile Phe Lys Ile Glu Arg Glu Gly Glu Cys Gln Arg Tyr
        835                 840                 845

Lys Pro Phe Lys Gln Leu His Asn Arg Arg Leu Leu Trp His Gly Ser
    850                 855                 860

Arg Thr Thr Asn Phe Ala Gly Ile Leu Ser Gln Gly Leu Arg Ile Ala
865                 870                 875                 880

Pro Pro Glu Ala Pro Val Thr Gly Tyr Met Phe Gly Lys Gly Ile Tyr
                885                 890                 895

Phe Ala Asp Met Val Ser Lys Ser Ala Asn Tyr Tyr His Thr Ser Gln
            900                 905                 910

Gly Asp Pro Ile Gly Leu Ile Leu Leu Gly Glu Val Ala Leu Gly Asn
        915                 920                 925

Met Tyr Glu Leu Lys His Ala Ser His Ile Ser Arg Leu Pro Lys Gly
    930                 935                 940

Lys His Ser Val Lys Gly Leu Gly Lys Thr Thr Pro Asp Pro Ser Ala
945                 950                 955                 960

Asn Ile Ser Leu Asp Gly Val Asp Val Pro Leu Gly Thr Gly Ile Ser
                965                 970                 975

Ser Gly Val Ile Asp Thr Ser Leu Leu Tyr Asn Glu Tyr Ile Val Tyr
            980                 985                 990
```

Asp Ile Ala Gln Val Asn Leu Lys Tyr Leu Leu Lys Leu Lys Phe Asn
        995   1000   1005

Phe Lys Thr Ser Leu Trp
    1010

<210> 138
<211> 5482
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (474)..(4016)

<400> 138
aaacaatgca atatttcgcg gagcagtgaa ttaatccgga aataatcgtc cgtgcccaga 60

gctttggagg ccaagtacca gggagctagt cccagagttg gtcgcagtct cagtaaaaac 120

gaatcgtagc caaccgcgat ccttgcaacc gtgctgtgtc gaaccaaaga aatcctattg 180

attttgggtc tgcaattgtg cattaaatat taagcaaaaa cgagggctgg tcgcgtggca 240

gccagtggca aattgttgct cctgcggcat aggcaggaca cctggataca ggatgcgggc 300

aagccagcga cggacaacgg cgaggcttgt gtaggacggg cagagcaact gctggaggag 360

agaactggac tgggagtgga aaacccgaaa gcccactgaa tattgcgctt gtttttttgtt 420

gcctattttt ttcggggcgt gtgtgtgcca aagcgtagca aacaagcaca aca atg   476
                                                           Met
                                                           1

gcc aac agc agc cga agt cgg gcc att ttg agc gtt aat ctc gat gcg   524
Ala Asn Ser Ser Arg Ser Arg Ala Ile Leu Ser Val Asn Leu Asp Ala
            5   10   15

gtc atg gcc aac gat ccg ctg agg gag ctc tcc gag gcc tgc aaa acg   572
Val Met Ala Asn Asp Pro Leu Arg Glu Leu Ser Glu Ala Cys Lys Thr
        20   25   30

ggc gag atc gcc aag gtg aag aag cta ata acg cct cag acc gtg aac   620
Gly Glu Ile Ala Lys Val Lys Lys Leu Ile Thr Pro Gln Thr Val Asn
    35   40   45

gcc agg gat acg gcg gga cgc aaa tcc aca cca ttg cat ttc gca gcg   668
Ala Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe Ala Ala
50   55   60   65

ggt tat gga cgc cgg gaa gtg gtt gaa ttc ctg ctg aac agc ggc gcc   716
Gly Tyr Gly Arg Arg Glu Val Val Glu Phe Leu Leu Asn Ser Gly Ala
                70   75   80

tcc ata cag gcg tgt gac gag ggt ggg ctg cac ccg ctg cac aac tgt   764
Ser Ile Gln Ala Cys Asp Glu Gly Gly Leu His Pro Leu His Asn Cys
            85   90   95

tgc tcc ttt ggc cac gcc gag gta gtt cga ttg ttg ctg aag gca ggt   812
Cys Ser Phe Gly His Ala Glu Val Val Arg Leu Leu Leu Lys Ala Gly
        100   105   110

gcc agt cca aac acc acc gac aac tgg aac tac acg cca ttg cac gag   860
Ala Ser Pro Asn Thr Thr Asp Asn Trp Asn Tyr Thr Pro Leu His Glu
    115   120   125

```
gcg gcc agc aag ggc aag gtg gat gtg tgc ctg gct ctg ttg cag cat     908
Ala Ala Ser Lys Gly Lys Val Asp Val Cys Leu Ala Leu Leu Gln His
130                 135                 140                 145

ggc gca aac cat acg atc cgc aac tcg gag cag aag aca cca ctg gag     956
Gly Ala Asn His Thr Ile Arg Asn Ser Glu Gln Lys Thr Pro Leu Glu
                150                 155                 160

ctg gcg gac gag gcg acg cgt ccc gta ttg acc ggc gaa tat cga aag     1004
Leu Ala Asp Glu Ala Thr Arg Pro Val Leu Thr Gly Glu Tyr Arg Lys
            165                 170                 175

gat gag ctg ctt gaa gcc gca cgc tcg ggg gcc gag gat cgc ctg ctg     1052
Asp Glu Leu Leu Glu Ala Ala Arg Ser Gly Ala Glu Asp Arg Leu Leu
        180                 185                 190

gcc cta ctc acg cca ctc aat gtc aac tgt cat gcc agc gat gga cga     1100
Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly Arg
    195                 200                 205

cgc tca acg ccg ctc cat ctg gca gcg ggc tac aat cgg atc ggc atc     1148
Arg Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Ile Gly Ile
210                 215                 220                 225

gtg gaa att ctg ctg gcc aac gga gcg gat gta cat gct aag gac aag     1196
Val Glu Ile Leu Leu Ala Asn Gly Ala Asp Val His Ala Lys Asp Lys
                230                 235                 240

ggc ggt ctg gtg ccg ctg cac aat gcc tgc tcc tac gga cac ttc gat     1244
Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Phe Asp
            245                 250                 255

gtg acc aag ctg ctt atc cag gcg ggc gcc aat gtc aac gcc aac gat     1292
Val Thr Lys Leu Leu Ile Gln Ala Gly Ala Asn Val Asn Ala Asn Asp
        260                 265                 270

ctg tgg gcc ttt acg ccg ctc cac gag gcc gcc tcc aaa agt cgc gtc     1340
Leu Trp Ala Phe Thr Pro Leu His Glu Ala Ala Ser Lys Ser Arg Val
    275                 280                 285

gag gtc tgc agc ctg ctg ctc agt cgt gga gcg gat ccc acc ctc cta     1388
Glu Val Cys Ser Leu Leu Leu Ser Arg Gly Ala Asp Pro Thr Leu Leu
290                 295                 300                 305

aac tgc cac agc aag tcg gcc atc gat gcg gcg ccc acc agg gag ctg     1436
Asn Cys His Ser Lys Ser Ala Ile Asp Ala Ala Pro Thr Arg Glu Leu
                310                 315                 320

aga gag cgg att gcc ttt gaa tac aag ggt cac tgc ctg ctg gac gcc     1484
Arg Glu Arg Ile Ala Phe Glu Tyr Lys Gly His Cys Leu Leu Asp Ala
            325                 330                 335

tgt cga aag tgt gat gtg tcc cgt gcc aag aag ctg gta tgc gca gag     1532
Cys Arg Lys Cys Asp Val Ser Arg Ala Lys Lys Leu Val Cys Ala Glu
        340                 345                 350

att gtt aac ttc gtg cat cca tat aca gga gac act ccg ctc cac ctg     1580
Ile Val Asn Phe Val His Pro Tyr Thr Gly Asp Thr Pro Leu His Leu
    355                 360                 365

gcc gtt gtc agt ccg gat ggg aag cgc aag cag ctg atg gaa ctg ctg     1628
Ala Val Val Ser Pro Asp Gly Lys Arg Lys Gln Leu Met Glu Leu Leu
370                 375                 380                 385

acc aga aag gga tcc ttg ctg aac gag aaa aac aag gct ttc ctc acg     1676
Thr Arg Lys Gly Ser Leu Leu Asn Glu Lys Asn Lys Ala Phe Leu Thr
                390                 395                 400
```

```
ccc ctg cat ttg gct gcc gag ctg ctt cac tac gat gcc atg gag gtg     1724
Pro Leu His Leu Ala Ala Glu Leu Leu His Tyr Asp Ala Met Glu Val
            405                 410                 415

ctg cta aag cag ggc gcc aag gtt aat gca ttg gac agt ctt gga caa     1772
Leu Leu Lys Gln Gly Ala Lys Val Asn Ala Leu Asp Ser Leu Gly Gln
        420                 425                 430

acg cca ctg cat cgg tgc gcc cgt gat gag caa gcg gtg cga ctg ctg     1820
Thr Pro Leu His Arg Cys Ala Arg Asp Glu Gln Ala Val Arg Leu Leu
    435                 440                 445

ctc tcg tac gca gcg gac acg aat atc gtt tcc ctt gag gga ctt acg     1868
Leu Ser Tyr Ala Ala Asp Thr Asn Ile Val Ser Leu Glu Gly Leu Thr
450                 455                 460                 465

gcc gct caa ttg gcc tcg gac agc gtg ctg aag ctg ctc aag aat cct     1916
Ala Ala Gln Leu Ala Ser Asp Ser Val Leu Lys Leu Leu Lys Asn Pro
                470                 475                 480

ccg gac agt gag aca cat tta ctg gag gca gcc aag gcg gga gat ctg     1964
Pro Asp Ser Glu Thr His Leu Leu Glu Ala Ala Lys Ala Gly Asp Leu
            485                 490                 495

gac act gtg cgc cgt ata gtg ctc aac aat ccg att tcg gtc aat tgc     2012
Asp Thr Val Arg Arg Ile Val Leu Asn Asn Pro Ile Ser Val Asn Cys
        500                 505                 510

cgg gat ttg gac gga cga cat tcc aca cct ttg cac ttt gct gct ggg     2060
Arg Asp Leu Asp Gly Arg His Ser Thr Pro Leu His Phe Ala Ala Gly
    515                 520                 525

ttt aat aga gtg cca gtg gtt cag ttt ctt ttg gaa cac ggc gcc gag     2108
Phe Asn Arg Val Pro Val Val Gln Phe Leu Leu Glu His Gly Ala Glu
530                 535                 540                 545

gtt tat gcg gct gac aag ggc gga ctg gtg ccc ctg cac aat gcc tgc     2156
Val Tyr Ala Ala Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala Cys
                550                 555                 560

tct tat ggg cac tat gag gta acc gaa ctg ctg gtc aag cac gga gcc     2204
Ser Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly Ala
            565                 570                 575

aat gta aat gta tcg gat ttg tgg aag ttt act cct ctt cat gaa gct     2252
Asn Val Asn Val Ser Asp Leu Trp Lys Phe Thr Pro Leu His Glu Ala
        580                 585                 590

gcc gcc aag gga aag tat gat att tgc aag ctg ctc ttg aaa cat ggc     2300
Ala Ala Lys Gly Lys Tyr Asp Ile Cys Lys Leu Leu Leu Lys His Gly
    595                 600                 605

gct gat cca atg aag aag aat cgg gat ggc gcg aca cca gcg gat ttg     2348
Ala Asp Pro Met Lys Lys Asn Arg Asp Gly Ala Thr Pro Ala Asp Leu
610                 615                 620                 625

gtt aag gaa tct gat cac gat gtt gca gag ctg ctg aga gga ccg tcc     2396
Val Lys Glu Ser Asp His Asp Val Ala Glu Leu Leu Arg Gly Pro Ser
                630                 635                 640

gct ctg cta gac gca gca aag aaa gga aac ttg gca cgg gta cag cga     2444
Ala Leu Leu Asp Ala Ala Lys Lys Gly Asn Leu Ala Arg Val Gln Arg
            645                 650                 655

ttg gtt aca ccg gaa tcc att aat tgc cgg gac gcg cag ggc agg aat     2492
Leu Val Thr Pro Glu Ser Ile Asn Cys Arg Asp Ala Gln Gly Arg Asn
        660                 665                 670
```

```
tcc aca cca ctt cac ctg gcc gcc gga tat aac aac ttt gag tgt gcc      2540
Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Phe Glu Cys Ala
    675                 680                 685

gag tac ctt ctg gag aat gga gcc gat gtt aat gca cag gac aag ggg      2588
Glu Tyr Leu Leu Glu Asn Gly Ala Asp Val Asn Ala Gln Asp Lys Gly
690                 695                 700                 705

gga cta ata cct ctg cac aat gcc agc agc tat ggg cat ttg gat att      2636
Gly Leu Ile Pro Leu His Asn Ala Ser Ser Tyr Gly His Leu Asp Ile
                710                 715                 720

gcg gca ctg cta att aag cac aag acg gtt gtc aat gcg aca gat aaa      2684
Ala Ala Leu Leu Ile Lys His Lys Thr Val Val Asn Ala Thr Asp Lys
            725                 730                 735

tgg gga ttc aca ccg ctc cac gag gct gca cag aag ggg cgc act caa      2732
Trp Gly Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr Gln
        740                 745                 750

ttg tgc tcg ctc ttg ttg gcc cac ggt gcc gat gcc tat atg aaa aac      2780
Leu Cys Ser Leu Leu Leu Ala His Gly Ala Asp Ala Tyr Met Lys Asn
    755                 760                 765

cag gag ggg cag acg ccc att gag ttg gcc acg gca gat gat gtt aag      2828
Gln Glu Gly Gln Thr Pro Ile Glu Leu Ala Thr Ala Asp Asp Val Lys
770                 775                 780                 785

tgc ttg ctc cag gac gcg atg gcc acc tcg ttg agt caa cag gcg ttg      2876
Cys Leu Leu Gln Asp Ala Met Ala Thr Ser Leu Ser Gln Gln Ala Leu
                790                 795                 800

agt gct tcc acg caa tcg ctg aca agc agt tcc ccg gca cca gat gca      2924
Ser Ala Ser Thr Gln Ser Leu Thr Ser Ser Ser Pro Ala Pro Asp Ala
            805                 810                 815

act gct gct gcg gct ccg ggc aca tct tca tcg tcc tca tcc gca atc      2972
Thr Ala Ala Ala Ala Pro Gly Thr Ser Ser Ser Ser Ser Ser Ala Ile
        820                 825                 830

cta tcg ccc acc acg gaa acg gtg ttg ctg ccc acc ggt gcc tcc atg      3020
Leu Ser Pro Thr Thr Glu Thr Val Leu Leu Pro Thr Gly Ala Ser Met
    835                 840                 845

att ctg agt gtt cct gtt cca ctt cca ctg tcc agt agc acg cgc atc      3068
Ile Leu Ser Val Pro Val Pro Leu Pro Leu Ser Ser Ser Thr Arg Ile
850                 855                 860                 865

agt ccc gcc caa gga gca gag gcc aat ggg gct gag ggc tcc tct tcg      3116
Ser Pro Ala Gln Gly Ala Glu Ala Asn Gly Ala Glu Gly Ser Ser Ser
                870                 875                 880

gat gat cta ctg ccg gat gcg gat acc ata aca aat gtg tcc gga ttc      3164
Asp Asp Leu Leu Pro Asp Ala Asp Thr Ile Thr Asn Val Ser Gly Phe
            885                 890                 895

cta agc agc cag cag ctg cat cat cta atc gaa ctg ttc gag cgc gaa      3212
Leu Ser Ser Gln Gln Leu His His Leu Ile Glu Leu Phe Glu Arg Glu
        900                 905                 910

caa atc acc ttg gac att cta gcc gag atg ggc cac gac gat ctc aag      3260
Gln Ile Thr Leu Asp Ile Leu Ala Glu Met Gly His Asp Asp Leu Lys
    915                 920                 925

cag gtg ggc gtc tcc gcc tac ggc ttc cgc cac aag ata ctc aag gga      3308
Gln Val Gly Val Ser Ala Tyr Gly Phe Arg His Lys Ile Leu Lys Gly
930                 935                 940                 945
```

```
atc gcc cag ctg agg tcc acc aca ggc att ggt aac aac gtg aat cta    3356
Ile Ala Gln Leu Arg Ser Thr Thr Gly Ile Gly Asn Asn Val Asn Leu
                950                 955                 960

tgc aca ttg ttg gtg gac ttg ctg ccg gac gat aag gag ttt gtg gcc    3404
Cys Thr Leu Leu Val Asp Leu Leu Pro Asp Asp Lys Glu Phe Val Ala
            965                 970                 975

gtc gag gag gag atg cag gcc acg att cgt gaa cat cgt gat aat gga    3452
Val Glu Glu Glu Met Gln Ala Thr Ile Arg Glu His Arg Asp Asn Gly
        980                 985                 990

cag gct gga ggt tat ttc act cga tat aac atc att cgg gtg caa aag    3500
Gln Ala Gly Gly Tyr Phe Thr Arg Tyr Asn Ile Ile Arg Val Gln Lys
    995                 1000                1005

gta caa aat cga aag ctg tgg gag cgt tat gct cat cga cgg caa gag    3548
Val Gln Asn Arg Lys Leu Trp Glu Arg Tyr Ala His Arg Arg Gln Glu
1010                1015                1020                1025

atc gcc gag gag aat ttc ctg cag tcc aac gag cgt atg ctc ttc cac    3596
Ile Ala Glu Glu Asn Phe Leu Gln Ser Asn Glu Arg Met Leu Phe His
                1030                1035                1040

ggt agt ccc ttc atc aac gca att gtg caa cgc gga ttc gac gag cgc    3644
Gly Ser Pro Phe Ile Asn Ala Ile Val Gln Arg Gly Phe Asp Glu Arg
            1045                1050                1055

cac gcc tac att ggc ggc atg ttt ggg gct ggc att tat ttc gcc gag    3692
His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu
        1060                1065                1070

cat agc tcg aaa agc aac cag tat gtg tac gga att ggc ggc ggc att    3740
His Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Ile
    1075                1080                1085

ggc tgt ccc tcg cac aag gat aag tcc tgc tac gtg tgt cct aga caa    3788
Gly Cys Pro Ser His Lys Asp Lys Ser Cys Tyr Val Cys Pro Arg Gln
1090                1095                1100                1105

ttg ctg ctg tgc cga gtg gcg tta ggc aaa tcc ttc ttg caa tac agt    3836
Leu Leu Leu Cys Arg Val Ala Leu Gly Lys Ser Phe Leu Gln Tyr Ser
                1110                1115                1120

gca atg aag atg gcc cat gca ccg ccg gga cac cac tcg gtg gtg ggc    3884
Ala Met Lys Met Ala His Ala Pro Pro Gly His His Ser Val Val Gly
            1125                1130                1135

aga ccc tcg gcg ggt ggc ttg cat ttc gcc gaa tac gtt gtc tat cgg    3932
Arg Pro Ser Ala Gly Gly Leu His Phe Ala Glu Tyr Val Val Tyr Arg
        1140                1145                1150

ggc gaa cag tct tat ccg gag tac ttg ata acc tac caa atc gtc aag    3980
Gly Glu Gln Ser Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Val Lys
    1155                1160                1165

ccc gat gac agc agt agt gga acg gag gat aca aga tgatggatgc         4026
Pro Asp Asp Ser Ser Ser Gly Thr Glu Asp Thr Arg
1170                1175                1180

cctctgtcgg gtccacgccc acaaccacgt cgcccgcgct gcaccagccg caaacgcaac 4086

aacaaccgca gcagcaacag cagcagcagc cgcaaccaca acaacagcag aaggcaccac 4146

tgccgttgcc accgccacaa cagcagacct cagctccagt tgccaagagg cggccgaaac 4206

atgccaaacc atcgctgcag ttgcagtatc agccctatca gccccagcac caccccggttg 4266

ttgcaaccgc cgctgctgtg accaccaccc aaccttcgcc cgctggcgtt tttgcgcaca 4326
```

gcaataacaa caataatacg agcagcggaa atgtgaataa taacaacaat gacatgtcgc 4386

cggtgtcgaa cagcaatagc tactcctcgg tggacaccaa ccagacgctg ctcaactcgc 4446

tggccaacca gcagcgcaac catcgacagc cacagaatca tcatcatcag cagcagcagc 4506

aggcgaatcg cagccaaaag tatagtcaat ttatgatcat cacacccgcc gtttccatag 4566

atcgcgactt cgagtacgag tcgcatttgg actttgagga tttcgccaat gcggcccaca 4626

acaatggcaa tctgtttcga cttggattgc ggcggagtga tagcagcagc gacgacagcg 4686

gccacagcag cgatagcagc agtttccgct cgaattacaa tccctacttg catcacagtc 4746

gccagcattt gctatcgaaa ggtggtaatg gtggtggcgg cggtgctagt cgtcacttct 4806

acgccttcac ctcgtcgtgg cgctggtgca gtcttctgtg cgccgccatg cgctgctttg 4866

gggccggggg agccgggcac gggaatgctc cgtacagcag ctcgctgcaa catcatcgac 4926

taagacgctg ttcgtcgtac aatgcggaga atgcctacga gcactttgcg gccccccttca 4986

aggcgcgcaa gacgcgcgac cacatgaata acatcaccta cgagttgtga cggtggtacg 5046

acatgctggt ggttattgat ctctatgtct ccgttggcct cccgtctatt ttattacaat 5106

tactagctat agatgtcgtg tcctgtgtgt ctctctctct ctcgttgttt attatattac 5166

tctataatat atcacgtaag ggcgagctag cgagatggat tcgttcggtt tggatcgaat 5226

tggattggat tcgattgggt taatcaaaag tgaagcacag tttttgagtg attttaattc 5286

gaaatacgga aaatgcgatt cgattatacg aggttacaag ttctttgccg atgaatgcat 5346

tacattacat tacattacgc tcgcgcgttt atttaagtgt ttaagcttag ttaatttaaa 5406

caataattaa aactccaatt aaatttaaat atacaaatac atatacatca atcgaaaaaa 5466

aaaaaaaaaa aaaaaa 5482

<210> 139
<211> 1181
<212> PRT
<213> Drosophila melanogaster

<400> 139
Met Ala Asn Ser Ser Arg Ser Arg Ala Ile Leu Ser Val Asn Leu Asp
1   5   10   15

Ala Val Met Ala Asn Asp Pro Leu Arg Glu Leu Ser Glu Ala Cys Lys
            20   25   30

Thr Gly Glu Ile Ala Lys Val Lys Lys Leu Ile Thr Pro Gln Thr Val
        35   40   45

Asn Ala Arg Asp Thr Ala Gly Arg Lys Ser Thr Pro Leu His Phe Ala
    50   55   60

Ala Gly Tyr Gly Arg Arg Glu Val Val Glu Phe Leu Leu Asn Ser Gly
65   70   75   80

Ala Ser Ile Gln Ala Cys Asp Glu Gly Gly Leu His Pro Leu His Asn
                85   90   95

Cys Cys Ser Phe Gly His Ala Glu Val Val Arg Leu Leu Leu Lys Ala
            100   105   110

Gly Ala Ser Pro Asn Thr Thr Asp Asn Trp Asn Tyr Thr Pro Leu His
        115   120   125

Glu Ala Ala Ser Lys Gly Lys Val Asp Val Cys Leu Ala Leu Leu Gln
130 135 140

His Gly Ala Asn His Thr Ile Arg Asn Ser Glu Gln Lys Thr Pro Leu
145 150 155 160

Glu Leu Ala Asp Glu Ala Thr Arg Pro Val Leu Thr Gly Glu Tyr Arg
165 170 175

Lys Asp Glu Leu Leu Glu Ala Ala Arg Ser Gly Ala Glu Asp Arg Leu
180 185 190

Leu Ala Leu Leu Thr Pro Leu Asn Val Asn Cys His Ala Ser Asp Gly
195 200 205

Arg Arg Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Arg Ile Gly
210 215 220

Ile Val Glu Ile Leu Leu Ala Asn Gly Ala Asp Val His Ala Lys Asp
225 230 235 240

Lys Gly Gly Leu Val Pro Leu His Asn Ala Cys Ser Tyr Gly His Phe
245 250 255

Asp Val Thr Lys Leu Leu Ile Gln Ala Gly Ala Asn Val Asn Ala Asn
260 265 270

Asp Leu Trp Ala Phe Thr Pro Leu His Glu Ala Ala Ser Lys Ser Arg
275 280 285

Val Glu Val Cys Ser Leu Leu Leu Ser Arg Gly Ala Asp Pro Thr Leu
290 295 300

Leu Asn Cys His Ser Lys Ser Ala Ile Asp Ala Ala Pro Thr Arg Glu
305 310 315 320

Leu Arg Glu Arg Ile Ala Phe Glu Tyr Lys Gly His Cys Leu Leu Asp
325 330 335

Ala Cys Arg Lys Cys Asp Val Ser Arg Ala Lys Lys Leu Val Cys Ala
340 345 350

Glu Ile Val Asn Phe Val His Pro Tyr Thr Gly Asp Thr Pro Leu His
355 360 365

Leu Ala Val Val Ser Pro Asp Gly Lys Arg Lys Gln Leu Met Glu Leu
370 375 380

Leu Thr Arg Lys Gly Ser Leu Leu Asn Glu Lys Asn Lys Ala Phe Leu
385 390 395 400

Thr Pro Leu His Leu Ala Ala Glu Leu Leu His Tyr Asp Ala Met Glu
405 410 415

Val Leu Leu Lys Gln Gly Ala Lys Val Asn Ala Leu Asp Ser Leu Gly
420 425 430

Gln Thr Pro Leu His Arg Cys Ala Arg Asp Glu Gln Ala Val Arg Leu
435 440 445

Leu Leu Ser Tyr Ala Ala Asp Thr Asn Ile Val Ser Leu Glu Gly Leu
450 455 460

Thr Ala Ala Gln Leu Ala Ser Asp Ser Val Leu Lys Leu Leu Lys Asn
465 470 475 480

Pro Pro Asp Ser Glu Thr His Leu Leu Glu Ala Ala Lys Ala Gly Asp
485 490 495

Leu Asp Thr Val Arg Arg Ile Val Leu Asn Asn Pro Ile Ser Val Asn
500 505 510

Cys Arg Asp Leu Asp Gly Arg His Ser Thr Pro Leu His Phe Ala Ala
515 520 525

Gly Phe Asn Arg Val Pro Val Val Gln Phe Leu Leu Glu His Gly Ala
530 535 540

Glu Val Tyr Ala Ala Asp Lys Gly Gly Leu Val Pro Leu His Asn Ala
545 550 555 560

Cys Ser Tyr Gly His Tyr Glu Val Thr Glu Leu Leu Val Lys His Gly
565 570 575

Ala Asn Val Asn Val Ser Asp Leu Trp Lys Phe Thr Pro Leu His Glu
580 585 590

Ala Ala Ala Lys Gly Lys Tyr Asp Ile Cys Lys Leu Leu Leu Lys His
595 600 605

Gly Ala Asp Pro Met Lys Lys Asn Arg Asp Gly Ala Thr Pro Ala Asp
610 615 620

Leu Val Lys Glu Ser Asp His Asp Val Ala Glu Leu Leu Arg Gly Pro
625 630 635 640

Ser Ala Leu Leu Asp Ala Ala Lys Lys Gly Asn Leu Ala Arg Val Gln
645 650 655

Arg Leu Val Thr Pro Glu Ser Ile Asn Cys Arg Asp Ala Gln Gly Arg
660 665 670

Asn Ser Thr Pro Leu His Leu Ala Ala Gly Tyr Asn Asn Phe Glu Cys
675 680 685

Ala Glu Tyr Leu Leu Glu Asn Gly Ala Asp Val Asn Ala Gln Asp Lys
690 695 700

Gly Gly Leu Ile Pro Leu His Asn Ala Ser Ser Tyr Gly His Leu Asp
705 710 715 720

Ile Ala Ala Leu Leu Ile Lys His Lys Thr Val Val Asn Ala Thr Asp
725 730 735

Lys Trp Gly Phe Thr Pro Leu His Glu Ala Ala Gln Lys Gly Arg Thr
740 745 750

Gln Leu Cys Ser Leu Leu Leu Ala His Gly Ala Asp Ala Tyr Met Lys
755 760 765

Asn Gln Glu Gly Gln Thr Pro Ile Glu Leu Ala Thr Ala Asp Asp Val
770 775 780

Lys Cys Leu Leu Gln Asp Ala Met Ala Thr Ser Leu Ser Gln Gln Ala
785 790 795 800

Leu Ser Ala Ser Thr Gln Ser Leu Thr Ser Ser Ser Pro Ala Pro Asp
805 810 815

Ala Thr Ala Ala Ala Ala Pro Gly Thr Ser Ser Ser Ser Ser Ser Ala
820 825 830

Ile Leu Ser Pro Thr Thr Glu Thr Val Leu Leu Pro Thr Gly Ala Ser
835 840 845

Met Ile Leu Ser Val Pro Val Pro Leu Pro Leu Ser Ser Ser Thr Arg
850 855 860

```
Ile Ser Pro Ala Gln Gly Ala Glu Ala Asn Gly Ala Glu Gly Ser Ser
865                 870                 875                 880

Ser Asp Asp Leu Leu Pro Asp Ala Asp Thr Ile Thr Asn Val Ser Gly
                885                 890                 895

Phe Leu Ser Ser Gln Gln Leu His His Leu Ile Glu Leu Phe Glu Arg
            900                 905                 910

Glu Gln Ile Thr Leu Asp Ile Leu Ala Glu Met Gly His Asp Asp Leu
        915                 920                 925

Lys Gln Val Gly Val Ser Ala Tyr Gly Phe Arg His Lys Ile Leu Lys
    930                 935                 940

Gly Ile Ala Gln Leu Arg Ser Thr Thr Gly Ile Gly Asn Asn Val Asn
945                 950                 955                 960

Leu Cys Thr Leu Leu Val Asp Leu Leu Pro Asp Asp Lys Glu Phe Val
                965                 970                 975

Ala Val Glu Glu Glu Met Gln Ala Thr Ile Arg Glu His Arg Asp Asn
            980                 985                 990

Gly Gln Ala Gly Gly Tyr Phe Thr Arg Tyr Asn Ile Ile Arg Val Gln
        995                 1000                1005

Lys Val Gln Asn Arg Lys Leu Trp Glu Arg Tyr Ala His Arg Arg Gln
    1010                1015                1020

Glu Ile Ala Glu Glu Asn Phe Leu Gln Ser Asn Glu Arg Met Leu Phe
1025                1030                1035                1040

His Gly Ser Pro Phe Ile Asn Ala Ile Val Gln Arg Gly Phe Asp Glu
                1045                1050                1055

Arg His Ala Tyr Ile Gly Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala
            1060                1065                1070

Glu His Ser Ser Lys Ser Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly
        1075                1080                1085

Ile Gly Cys Pro Ser His Lys Asp Lys Ser Cys Tyr Val Cys Pro Arg
    1090                1095                1100

Gln Leu Leu Leu Cys Arg Val Ala Leu Gly Lys Ser Phe Leu Gln Tyr
1105                1110                1115                1120

Ser Ala Met Lys Met Ala His Ala Pro Pro Gly His His Ser Val Val
                1125                1130                1135

Gly Arg Pro Ser Ala Gly Gly Leu His Phe Ala Glu Tyr Val Val Tyr
            1140                1145                1150

Arg Gly Glu Gln Ser Tyr Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Val
        1155                1160                1165

Lys Pro Asp Asp Ser Ser Ser Gly Thr Glu Asp Thr Arg
    1170                1175                1180
```

<210> 140
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 140
ggcctgaagg tatggtcgat 20

<210> 141
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 141
tgagggcatt acagtttgtt 20

<210> 142
<211> 346
<212> DNA
<213> Homo sapiens

<400> 142
ggcctgaagg tatggtcgat ggataaatag ttattttaag aaactaattc cactgaacct 60
aaaatcatca aagcagcagt ggcctctacg ttttactcct ttgctgaaaa aaaatcatct 120
tgcccacagg cctgtggcaa aaggataaaa atgtgaacga agtttaacat tctgacttga 180
taaagcttta ataatgtaca gtgttttcta aatatttcct gtttttcag cactttaaca 240
gatgccattc caggttaaac tgggttgtct gtactaaatt ataaacagag ttaacttgaa 300
ccttttatat gttatgcatt gattctaaca aactgtaatg ccctca 346

<210> 143
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 143
gccgaattcg gcctgaaggt atggtcgat 29

<210> 144
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 144
gccgaattct agatgagggc attacagttt gtt 33

<210> 145
<211> 362
<212> DNA
<213> Homo sapiens

<400> 145
gaattcggcc tgaaggtatg gtcgatggat aaatagttat tttaagaaac taattccact 60
gaacctaaaa tcatcaaagc agcagtggcc tctacgtttt actcctttgc tgaaaaaaaa 120
tcatcttgcc cacaggcctg tggcaaaagg ataaaaatgt gaacgaagtt taacattctg 180
acttgataaa gctttaataa tgtacagtgt tttctaaata tttcctgttt tttcagcact 240
ttaacagatg ccattccagg ttaaactggg ttgtctgtac taaattataa acagagttaa 300
cttgaacctt ttatatgtta tgcattgatt ctaacaaact gtaatgccct catctagaat 360
tc 362

<210> 146

<211> 5616
<212> DNA
<213> Homo sapiens

<400> 146
tgaatgcctg ctggtgaagg ccagatcaga tttcaacctg ggactggatt acagaggatt 60
gtttctaata acaacatcaa tattctagaa gtccctgaca gcctagaaat aagctgtttg 120
tcttctataa agcattgcta tagtgatgaa tagtatgagt aactgataca tactcaactg 180
ctactgttcc ctttgaggaa atgtttacag gggcggcctt ttaacatatc tcaggctcat 240
tttcattgca attatccatt tctaaaacaa gattgcttcg atctagactt ggaaatggaa 300
aataagaaaa ccaatgcttt ttcaaatgtt cacaattcac acactacatt tgttttgtta 360
tgcatgacgt gtctataaca aatatacaca tacgacaggc aacaagcttg tttttgattt 420
gccagacatg catcattggc tattgtttgt ttgtttttttg ttttttttgtg tttttttgggt 480
tactttgaaa atgagccaga gccttcttga ggatattttg cacaaagtca cgctgacaaa 540
atcattagca gtgcaaccca agcttctggc tgagcaagat tcagtttcca ctttttaaaa 600
tttttttatt ttgctctgta gctgcacttc tcgttatcat aaattgagat gaaaaggaaa 660
aaacatcaag ttttagtacc tttttatgaa ttggcctatc ttacaagaga agggcacaaa 720
caccaacctg acttaggaac gcctaaattc agagaagtca aagccggtga aggccacttg 780
ctctttccaa cacaagcctg ccacagaggt cttcgggaca gtactggaga tgcaggttga 840
cacgggcttg agttccaagg tgaaaaaact ggggaggctg tgaaggaaga gctgcattaa 900
ggagggtgag gagcgtgtgg ttctgtatca tggcagcccc aatggatcca ggggatgcct 960
ccaaaaaata catgcttccc ttcccttaat ctgtactgtt gggattgtta cccctccaaa 1020
ttagctgcct tatttcaaaa gtcagtgaaa ttactgcact tgatgagggt cacaaaaata 1080
ccacttgatt gtttctttag ttgagaatgc tgggattcag actcgaatag tggatagata 1140
cacacaaatg caaggacttt tttgtttact ccagatttgg ggtttatttt gagtggcatg 1200
cttcaaatag ttcataaaga tccttgcatt aaatttctga accatttctt caaacttctt 1260
agtgtgttta gacaaggaga acaaaaattg aaaccaaagc cctttctgtt atttttttcaa 1320
tgaaggtgag aaagaaatac catacaattt tctttgtgaa attactgttt attttcatca 1380
acatttacca agtgccattg acatttataa aaaaaatgat cctttatagt tcttacactt 1440
gcccttttca ccttaactga atatgaattg agtgcactaa cttatttact tgatatactg 1500
tgcatctact ctgctttgaa gcgaaagaaa tataaacacg aggaggaata ggaaagacag 1560
tgtgacacaa acttgccatt gcaattcaaa gccctgaaaa cgatgggttt aatgcaaggt 1620
gattaagctg tgacctcctt taatctcctg aagcaaaata aaatggttac atgcaaaact 1680
tctagaaata cactcttaaa atatatacat tttgctttga ttttggcttc aacccagtgc 1740
tggaactagg catccagact agtttgaatg tttgtagctg aatttttatg ggtcctcaaa 1800
attaaatcga gaattagcct cagttgttgc ttcttttgaa gtttcagtga cccaagctgg 1860
gtgtttgtgt cttggctact tgtttaatag cactagaatt ccaggtgaag ctttgagagt 1920
tgatattcat taagagggct ttttttcccc ttctttcctt ctcttttgct gtaacaaagg 1980
gttgaagaaa ttgccatctg tgtagtttc agtagctgtc aagtgtgtct tacttacctt 2040
cccccagacg tagtttaaaa tggtaaacac agctgtgatt tttagttaag taaaagagtt 2100
aatatgatat agatatggaa agctttatgg cttcattaaa aagataaacc actacctaac 2160
tgtggttgta tgttgtttcc atcatactaa ctagatgaat ggatgcgcca gttttcatct 2220
tggtccttac acttgagaag ttaaactgtg gttcagtatt taaactgcca gtgttatacg 2280
tctcatgctc tgtgtgccag gtgaaggtac tgtgtaagga agacatttgc ggtgcttctt 2340
gtcctataat gattcaagta tatagtagtt cttgaaagag tgtgcatata ttactcatct 2400
gcttaagaga gtgggttaat ggatatatca gaggagccaa atacattttt ttcagaactt 2460
gaaaaccaaa ggtcatcatg agtgcactca aaagttagga caagtttatt acatttggga 2520
ttttcatctg tagccgtatg aagaaccctt tccaatataa aagcatggca ttaaattagg 2580
ctgaagtctt ttattttttg tatatgtact atatagaaat actagcaagt taggatcatc 2640
caatatggcc taccccgaaa tggcccctct gtttccctaa ccacatggaa gaaagaatct 2700
gaacgtctcc accggctcta cccgagttcc aaaactaaag ggcttctcca gacctgatgg 2760
ttccagttta cctgctgttg gcctgctgga tacttgactc aggcataaat taagtgccct 2820
ggtcccgaac tttctcccag tatttgacct ccttccctct ttcctaaatt actagtctgg 2880
aattaaaatt agctccagca atgacctttg actccattca ttttctcctc atcttgggtc 2940
ttaaaaaagg agaccagata cctcctagct tttgtatcac aaccaggaat gggtattagg 3000
cctcatgcgc tttgctcaga acactgccgc tttgttaaca aatgacagca tggaacccag 3060
agttttgatt cgatgcaaaa taacagcagt gcaaccagga ttcttgtttt cctttttcctt 3120
cttggagttt ggaatttcta gcttttcaag cagcataagt agaatcaaca ttaggatgtt 3180
ttcatgaaat agcatcctta tacttctttg agcttgatgt tagtggctag actgatttcc 3240
ctttgctctc aaaatacaaa gtgcattgaa gtatacagag aaatgcctga atatggcaag 3300
caaataatgt agattaacat tctattattg tatccgtttt acaaaaaata aaattttgat 3360
atatgccgga gaacggcatt agaatgcaat aagttgtcta ggtttttctg tttcagtgtc 3420
tctcccaatg gcacgaaggg ttattgggca ttgtccccac ccccgccttt ttaacatgtg 3480
cactatctgg attcctgtaa atggccttgc aaacagaagt ggtgtgtatt ttcaagcacc 3540
tttcccccat tgtatccgaa tccctcttgt gtgatatctg tgacaaatac cattcttctt 3600
gtgttttctg ttgggactaa ttgtctcacg taaagctata gaccttacta atttggcagg 3660
tattcaaaac tgccattaag ataggatttc atgtcagata cgtatttaaa gagtaaagtc 3720
aaatttgttt aatgtcagat cagtgacaga agtgaaaaga aagtaattgt gaaagtgatg 3780
tttgagctat tgtacacatc tagcatatgg aaagcaaatg cactcgaaaa ctactattct 3840
agaacatgag gcttcttcag caacttgtgc actctgccat taataaatta aatttttccc 3900

ctctagaaag ccttaactat ggcggaaact ttttaacctt ttatatttta ataaataaaa 3960
cattgtagtc ccatttctta gtgtttgaaa ggtgtgtcag tgagtcggcc atgtctccat 4020
gtgtttcaga cctgttcatc ttattttatg atggtatatt tcataagtaa tattccctta 4080
catgcaatgg agctgattaa aattaatcca tttcaatttc tccatattgg aacttcctca 4140
gctaccagat ttctggtttg gagaagtgct ggaaagattt caaagcctat tcagttgtgt 4200
atgtggggat acgacagcaa ctgtgatacc ttgtagaata tgagtgatat gcaagctgtg 4260
ttttttaatt gttttaaaat gtaaattatg gttatgctaa agtgaaaacc tagaggaagc 4320
taatgatttt atatactttg cacgaccaaa tatggtcgta gtatgacgag ttttatacat 4380
tgccagagag ttctgcctcc tctgaaataa cattcgcact gtagattgca tttcggcttt 4440
tcctcctttc acattctttt ttgctttaca cttcacgtct tcgcacctgc cctacctccc 4500
atcctttcaa agaggtttct ttcacgttcc agaattcaga ttgttctgtg atttctttta 4560
catcagtcta cccatttctg caggcagccc tgaaagccct tgtgttgatt cagagtgttt 4620
gcagagaaat gcagttgaac cctggtagtg gggtgtccct cacacacccg cgcacccctc 4680
ccaaagttca ggatgaaagg ctagaaaacc cattcaaagt taggaaagaa cacagatctt 4740
tgaggccgat agcctagacc tagaagatga ccttgagtat gtaaacattg tctccgtgac 4800
acaaaacact gaaactcttc atgtgcatat aacacctgct tctgctccca ttgtttcaag 4860
ctcatcttat ctttgtagta gtaatgtttg tctttgatac ctacaaacta aaaaggtact 4920
tttatcaagg tttctcaaaa catttacaaa accagctttg agaaaatgtt atgttgcctg 4980
gcaacagcac tcggagtagt aattgtgttt tctcattgtg atgttggtct gtgtgagcaa 5040
ccagtgtagt gactctttgg ttcattattc gtgttgtttt tatttttagt ctctgtgtga 5100
cccaacagtg gcaggggtta caaccccctc tcctttcttt tttgtattta tctatttgta 5160
ggattgtcag atcaagtaca agatgcccag ttaagtttga atttcagaga aacaatttca 5220
cgttaagaat gtttcatgca atatttggca tatatttaca gtaaaagcat tcattatttg 5280
tctgaaattc aaatttaact gagcatgctg gtttttctca ttgtttggtt tttctaaatc 5340
tggcaatcct acagctgtgg tcatgggaaa tcacctacag catgttaaag tcctctagtc 5400
atcatctcgt cacctgaaat ggaagtcctt tttccctcac cctccacttc tttccaaagg 5460
agggcatcaa ggaacttaac ctgcctgcct ggtgggtttc tatttaagac atctttgtga 5520
ttatatttaa cctgcaattg tgctttggct taatgtctag ctcactgtac ttgtaaatga 5580
ttaatattca ataaaaccat ttttaaagta aaaaaa 5616

<210> 147
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 147
gccgaattcc ttgtttttga tttgccaga 29

<210> 148
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 148
gccgaattcc ggctttgact tctctgaatt tagg 34

<210> 149
<211> 372
<212> DNA
<213> Homo sapiens

<400> 149
gaattccttg tttttgattt gccagacatg catcattggc tattgtttgt ttgttttttg 60
ttttttgtg tttttgggt tactttgaaa atgagccaga gccttcttga ggatattttg 120
cacaaagtca cgctgacaaa atcattagca gtgcaaccca agcttctggc tgagcaagat 180
tcagtttcca ctttttaaaa tttttttatt ttgctctgta gctgcacttc tcgttatcat 240
aaattgagat gaaaaggaaa aaacatcaag ttttagtacc tttttatgaa ttggcctatc 300
ttacaagaga agggcacaaa caccaacctg acttaggaac gcctaaattc agagaagtca 360
aagccggaat tc 372

<210> 150
<211> 1320
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1317)

<400> 150
```
atg gcg gag gat gtt tcc tca gcg gcc ccg agc ccg cgg cgg tgt gcg      48
Met Ala Glu Asp Val Ser Ser Ala Ala Pro Ser Pro Arg Arg Cys Ala
  1               5                  10                  15

gat ggt agg gat gcc gac cct act gag gag cag atg gca gaa aca gag      96
Asp Gly Arg Asp Ala Asp Pro Thr Glu Glu Gln Met Ala Glu Thr Glu
             20                  25                  30

aga aac gac gag gag cag ttc gaa tgc cag gaa ctg ctc gag tgc cag     144
Arg Asn Asp Glu Glu Gln Phe Glu Cys Gln Glu Leu Leu Glu Cys Gln
         35                  40                  45

gtg cag gtg ggg gcc ccc gag gag gag gag gag gag gag gag gac gcg     192
Val Gln Val Gly Ala Pro Glu Glu Glu Glu Glu Glu Glu Glu Asp Ala
     50                  55                  60

ggc ctg gtg gcc gag gcc gag gcc gtg gct gcc ggc tgg atg ctc gat     240
Gly Leu Val Ala Glu Ala Glu Ala Val Ala Ala Gly Trp Met Leu Asp
 65                  70                  75                  80

ttc ctc tgc ctc tct ctt tgc cga gct ttc cgc gac ggc cgc tcc gag     288
Phe Leu Cys Leu Ser Leu Cys Arg Ala Phe Arg Asp Gly Arg Ser Glu
                 85                  90                  95

gac ttc cgc agg acc cgc aac agc gca gag gct att att cat gga cta     336
Asp Phe Arg Arg Thr Arg Asn Ser Ala Glu Ala Ile Ile His Gly Leu
            100                 105                 110

tcc agt cta aca gct tgc cag ttg aga acg ata tac ata tgt cag ttt     384
Ser Ser Leu Thr Ala Cys Gln Leu Arg Thr Ile Tyr Ile Cys Gln Phe
        115                 120                 125

ttg aca aga att gca gca gga aaa acc ctt gat gca cag ttt gaa aat     432
Leu Thr Arg Ile Ala Ala Gly Lys Thr Leu Asp Ala Gln Phe Glu Asn
    130                 135                 140

gat gaa cga att aca ccc ttg gaa tca gcc ctg atg att tgg ggt tca     480
Asp Glu Arg Ile Thr Pro Leu Glu Ser Ala Leu Met Ile Trp Gly Ser
145                 150                 155                 160

att gaa aag gaa cat gac aaa ctt cat gaa gaa ata cag aat tta att     528
Ile Glu Lys Glu His Asp Lys Leu His Glu Glu Ile Gln Asn Leu Ile
                165                 170                 175

aaa att cag gct ata gct gtt tgt atg gaa aat ggc aac ttt aaa gaa     576
Lys Ile Gln Ala Ile Ala Val Cys Met Glu Asn Gly Asn Phe Lys Glu
            180                 185                 190

gca gaa gaa gtc ttt gaa aga ata ttt ggt gat cca aat tct cat atg     624
Ala Glu Glu Val Phe Glu Arg Ile Phe Gly Asp Pro Asn Ser His Met
        195                 200                 205

cct ttc aaa agc aaa ttg ctt atg ata atc tct cag aaa gat aca ttt     672
Pro Phe Lys Ser Lys Leu Leu Met Ile Ile Ser Gln Lys Asp Thr Phe
    210                 215                 220

cat tcc ttt ttt caa cac ttc agc tac aac cac atg atg gag aaa att     720
His Ser Phe Phe Gln His Phe Ser Tyr Asn His Met Met Glu Lys Ile
225                 230                 235                 240
```

aag agt tat gtg aat tat gtg cta agt gaa aaa tca tca acc ttt cta 768
Lys Ser Tyr Val Asn Tyr Val Leu Ser Glu Lys Ser Ser Thr Phe Leu
245 250 255

atg aag gca gcg gca aaa gta gta gaa agc aaa agg aca aga aca ata 816
Met Lys Ala Ala Ala Lys Val Val Glu Ser Lys Arg Thr Arg Thr Ile
260 265 270

act tct caa gat aaa cct agt ggt aat gat gtt gaa atg gaa act gaa 864
Thr Ser Gln Asp Lys Pro Ser Gly Asn Asp Val Glu Met Glu Thr Glu
275 280 285

gct aat ttg gat aca aga aaa agt gtt agt gac aaa cag tct gcg gta 912
Ala Asn Leu Asp Thr Arg Lys Ser Val Ser Asp Lys Gln Ser Ala Val
290 295 300

act gaa tcc tca gag ggt aca gta tcc tta ttg agg tct cac aag aat 960
Thr Glu Ser Ser Glu Gly Thr Val Ser Leu Leu Arg Ser His Lys Asn
305 310 315 320

ctt ttc tta tct aag ttg caa cat gga acc cag caa caa gac ctt aat 1008
Leu Phe Leu Ser Lys Leu Gln His Gly Thr Gln Gln Gln Asp Leu Asn
325 330 335

aag aaa gaa aga aga gta gga act cct caa agt aca aaa aag aaa aaa 1056
Lys Lys Glu Arg Arg Val Gly Thr Pro Gln Ser Thr Lys Lys Lys Lys
340 345 350

gaa agc aga aga gcc act gaa agc aga ata cct gtt tca aag agt cag 1104
Glu Ser Arg Arg Ala Thr Glu Ser Arg Ile Pro Val Ser Lys Ser Gln
355 360 365

ccg gta act cct gaa aaa cat cga gct aga aaa aga cag gca tgg ctt 1152
Pro Val Thr Pro Glu Lys His Arg Ala Arg Lys Arg Gln Ala Trp Leu
370 375 380

tgg gaa gaa gac aag aat ttg aga tct ggc gtg agg aaa tat gga gag 1200
Trp Glu Glu Asp Lys Asn Leu Arg Ser Gly Val Arg Lys Tyr Gly Glu
385 390 395 400

gga aac tgg tct aaa ata ctg ttg cat tat aaa ttc aac aac cgg aca 1248
Gly Asn Trp Ser Lys Ile Leu Leu His Tyr Lys Phe Asn Asn Arg Thr
405 410 415

agt gtc atg tta aaa gac aga tgg agg acc atg aag aaa cta aaa ctg 1296
Ser Val Met Leu Lys Asp Arg Trp Arg Thr Met Lys Lys Leu Lys Leu
420 425 430

att tcc tca gac agc gaa gac tga 1320
Ile Ser Ser Asp Ser Glu Asp
435

<210> 151
<211> 439
<212> PRT
<213> Homo sapiens

<400> 151
Met Ala Glu Asp Val Ser Ser Ala Ala Pro Ser Pro Arg Arg Cys Ala
1 5 10 15

Asp Gly Arg Asp Ala Asp Pro Thr Glu Glu Gln Met Ala Glu Thr Glu
20 25 30

Arg Asn Asp Glu Glu Gln Phe Glu Cys Gln Glu Leu Leu Glu Cys Gln
35 40 45

Val Gln Val Gly Ala Pro Glu Glu Glu Glu Glu Glu Glu Glu Asp Ala
50 55 60

Gly Leu Val Ala Glu Ala Glu Ala Val Ala Ala Gly Trp Met Leu Asp
65 70 75 80

Phe Leu Cys Leu Ser Leu Cys Arg Ala Phe Arg Asp Gly Arg Ser Glu
85 90 95

Asp Phe Arg Arg Thr Arg Asn Ser Ala Glu Ala Ile Ile His Gly Leu
100 105 110

Ser Ser Leu Thr Ala Cys Gln Leu Arg Thr Ile Tyr Ile Cys Gln Phe
115 120 125

Leu Thr Arg Ile Ala Ala Gly Lys Thr Leu Asp Ala Gln Phe Glu Asn
130 135 140

Asp Glu Arg Ile Thr Pro Leu Glu Ser Ala Leu Met Ile Trp Gly Ser
145 150 155 160

Ile Glu Lys Glu His Asp Lys Leu His Glu Glu Ile Gln Asn Leu Ile
165 170 175

Lys Ile Gln Ala Ile Ala Val Cys Met Glu Asn Gly Asn Phe Lys Glu
180 185 190

Ala Glu Glu Val Phe Glu Arg Ile Phe Gly Asp Pro Asn Ser His Met
195 200 205

Pro Phe Lys Ser Lys Leu Leu Met Ile Ile Ser Gln Lys Asp Thr Phe
210 215 220

His Ser Phe Phe Gln His Phe Ser Tyr Asn His Met Met Glu Lys Ile
225 230 235 240

Lys Ser Tyr Val Asn Tyr Val Leu Ser Glu Lys Ser Ser Thr Phe Leu
245 250 255

Met Lys Ala Ala Ala Lys Val Val Glu Ser Lys Arg Thr Arg Thr Ile
260 265 270

Thr Ser Gln Asp Lys Pro Ser Gly Asn Asp Val Glu Met Glu Thr Glu
275 280 285

Ala Asn Leu Asp Thr Arg Lys Ser Val Ser Asp Lys Gln Ser Ala Val
290 295 300

Thr Glu Ser Ser Glu Gly Thr Val Ser Leu Leu Arg Ser His Lys Asn
305 310 315 320

Leu Phe Leu Ser Lys Leu Gln His Gly Thr Gln Gln Gln Asp Leu Asn
325 330 335

Lys Lys Glu Arg Arg Val Gly Thr Pro Gln Ser Thr Lys Lys Lys Lys
340 345 350

Glu Ser Arg Arg Ala Thr Glu Ser Arg Ile Pro Val Ser Lys Ser Gln
355 360 365

Pro Val Thr Pro Glu Lys His Arg Ala Arg Lys Arg Gln Ala Trp Leu
370 375 380

Trp Glu Glu Asp Lys Asn Leu Arg Ser Gly Val Arg Lys Tyr Gly Glu
385 390 395 400

Gly Asn Trp Ser Lys Ile Leu Leu His Tyr Lys Phe Asn Asn Arg Thr
405 410 415

Ser Val Met Leu Lys Asp Arg Trp Arg Thr Met Lys Lys Leu Lys Leu
420 425 430

```
Ile Ser Ser Asp Ser Glu Asp
        435


<210> 152
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 152
gccccgggga tcctcatggc ggaggatgtt tcctcagcg                            39


<210> 153
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 153
tcccggggat cctcacacca ggcccgcgtc ctc                                  33


<210> 154
<211> 201
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(201)

<400> 154
atg gcg gag gat gtt tcc tca gcg gcc ccg agc ccg cgg ggc tgt gcg      48
Met Ala Glu Asp Val Ser Ser Ala Ala Pro Ser Pro Arg Gly Cys Ala
  1               5                  10                  15

gat ggt agg gat gcc gac cct act gag gag cag atg gca gaa aca gag      96
Asp Gly Arg Asp Ala Asp Pro Thr Glu Glu Gln Met Ala Glu Thr Glu
             20                  25                  30

aga aac gac gag gag cag ttc gaa tgc cag gaa ctg ctc gag tgc cag      144
Arg Asn Asp Glu Glu Gln Phe Glu Cys Gln Glu Leu Leu Glu Cys Gln
         35                  40                  45

gtg cag gtg ggg gcc ccc gag gag gag gag gag gag gag gag gac gcg      192
Val Gln Val Gly Ala Pro Glu Glu Glu Glu Glu Glu Glu Glu Asp Ala
     50                  55                  60

ggc ctg gtg                                                           201
Gly Leu Val
 65


<210> 155
<211> 67
<212> PRT
<213> Homo sapiens

<400> 155
Met Ala Glu Asp Val Ser Ser Ala Ala Pro Ser Pro Arg Gly Cys Ala
  1               5                  10                  15
```

```
Asp Gly Arg Asp Ala Asp Pro Thr Glu Glu Gln Met Ala Glu Thr Glu
             20                  25                  30

Arg Asn Asp Glu Glu Gln Phe Glu Cys Gln Glu Leu Leu Glu Cys Gln
         35                  40                  45

Val Gln Val Gly Ala Pro Glu Glu Glu Glu Glu Glu Glu Glu Asp Ala
     50                  55                  60

Gly Leu Val
 65


<210> 156
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 156
cgcaggatcc ccttcactcc tcttcatgag gcagcttc                         38


<210> 157
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 157
ggatccgcta aatatctgta tctccatctt taacaagatc caaaggag              48


<210> 158
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 158
gccgacttcg agtttgagca g                                           21


<210> 159
<211> 1103
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (9)..(1094)

<400> 159
ggatcccc ttc act cct ctt cat gag gca gct tct aag aac agg gtt gaa  50
         Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val Glu
           1               5                  10

gta tgt tct ctt ctc tta agt tat ggt gca gac cca aca ctg ctc aat   98
Val Cys Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr Leu Leu Asn
 15                  20                  25                  30
```

tgt cac aat aaa agt gct ata gac ttg gct ccc aca cca cag tta aaa 146
Cys His Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro Gln Leu Lys
35 40 45

gaa aga tta gca tat gaa ttt aaa ggc cac tcg ttg ctg caa gct gca 194
Glu Arg Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala Ala
50 55 60

cga gaa gct gat gtt act cga atc aaa aaa cat ctc tct ctg gaa atg 242
Arg Glu Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser Leu Glu Met
65 70 75

gtg aat ttc aag cat cct caa aca cat gaa aca gca ttg cat tgt gct 290
Val Asn Phe Lys His Pro Gln Thr His Glu Thr Ala Leu His Cys Ala
80 85 90

gct gca tct cca tat ccc aaa aga aag caa ata tgt gaa ctg ttg cta 338
Ala Ala Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu Leu Leu Leu
95 100 105 110

aga aaa gga gca aac atc aat gaa aag act aaa gaa ttc ttg act cct 386
Arg Lys Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe Leu Thr Pro
115 120 125

ctg cac gtg gca tct gag aaa gct cat aat gat gtt gtt gaa gta gtg 434
Leu His Val Ala Ser Glu Lys Ala His Asn Asp Val Val Glu Val Val
130 135 140

gtg aaa cat gaa gca aag gtt aat gct ctg gat aat ctt ggt cag act 482
Val Lys His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu Gly Gln Thr
145 150 155

tct cta cac aga gct gca tat tgt ggt cat cta caa acc tgc cgc cta 530
Ser Leu His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr Cys Arg Leu
160 165 170

ctc ctg agc tat ggg tgt gat cct aac att ata tcc ctt cag ggc ttt 578
Leu Leu Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe
175 180 185 190

act gct tta cag atg gga aat gaa aat gta cag caa ctc ctc caa gag 626
Thr Ala Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu Leu Gln Glu
195 200 205

ggt atc tca tta ggt aat tca gag gca gac aga caa ttg ctg gaa gct 674
Gly Ile Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala
210 215 220

gca aag gct gga gat gtc gaa act gta aaa aaa ctg tgt act gtt cag 722
Ala Lys Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys Thr Val Gln
225 230 235

agt gtc aac tgc aga gac att gaa ggg cgt cag tct aca cca ctt cat 770
Ser Val Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr Pro Leu His
240 245 250

ttt gca gct ggg tat aac aga gtg tcc gtg gtg gaa tat ctg cta cag 818
Phe Ala Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu Gln
255 260 265 270

cat gga gct gat gtg cat gct aaa gat aaa gga ggc ctt gta cct ttg 866
His Gly Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro Leu
275 280 285

cac aat gca tgt tct tat gga cat tat gaa gtt gca gaa ctt ctt gtt 914
His Asn Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu Val
290 295 300

```
aaa cat gga gca gta gtt aat gta gct gat tta tgg aaa ttt aca cct    962
Lys His Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys Phe Thr Pro
        305                 310                 315

tta cat gaa gca gca gca aaa gga aaa tat gaa att tgc aaa ctt ctg   1010
Leu His Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu
    320                 325                 330

ctc cag cat ggt gca gac cct aca aaa aaa aac agg gat gga aat act   1058
Leu Gln His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr
335                 340                 345                 350

cct ttg gat ctt gtt aaa gat gga gat aca gat att tagcggatc         1103
Pro Leu Asp Leu Val Lys Asp Gly Asp Thr Asp Ile
                355                 360


&lt;210&gt; 160
&lt;211&gt; 362
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 160
Phe Thr Pro Leu His Glu Ala Ala Ser Lys Asn Arg Val Glu Val Cys
  1               5                  10                  15

Ser Leu Leu Leu Ser Tyr Gly Ala Asp Pro Thr Leu Leu Asn Cys His
            20                  25                  30

Asn Lys Ser Ala Ile Asp Leu Ala Pro Thr Pro Gln Leu Lys Glu Arg
        35                  40                  45

Leu Ala Tyr Glu Phe Lys Gly His Ser Leu Leu Gln Ala Ala Arg Glu
    50                  55                  60

Ala Asp Val Thr Arg Ile Lys Lys His Leu Ser Leu Glu Met Val Asn
65                  70                  75                  80

Phe Lys His Pro Gln Thr His Glu Thr Ala Leu His Cys Ala Ala Ala
                85                  90                  95

Ser Pro Tyr Pro Lys Arg Lys Gln Ile Cys Glu Leu Leu Leu Arg Lys
            100                 105                 110

Gly Ala Asn Ile Asn Glu Lys Thr Lys Glu Phe Leu Thr Pro Leu His
        115                 120                 125

Val Ala Ser Glu Lys Ala His Asn Asp Val Val Glu Val Val Val Lys
    130                 135                 140

His Glu Ala Lys Val Asn Ala Leu Asp Asn Leu Gly Gln Thr Ser Leu
145                 150                 155                 160

His Arg Ala Ala Tyr Cys Gly His Leu Gln Thr Cys Arg Leu Leu Leu
                165                 170                 175

Ser Tyr Gly Cys Asp Pro Asn Ile Ile Ser Leu Gln Gly Phe Thr Ala
            180                 185                 190

Leu Gln Met Gly Asn Glu Asn Val Gln Gln Leu Leu Gln Glu Gly Ile
        195                 200                 205

Ser Leu Gly Asn Ser Glu Ala Asp Arg Gln Leu Leu Glu Ala Ala Lys
    210                 215                 220

Ala Gly Asp Val Glu Thr Val Lys Lys Leu Cys Thr Val Gln Ser Val
225                 230                 235                 240

Asn Cys Arg Asp Ile Glu Gly Arg Gln Ser Thr Pro Leu His Phe Ala
                245                 250                 255
```

Ala Gly Tyr Asn Arg Val Ser Val Val Glu Tyr Leu Leu Gln His Gly
260 265 270

Ala Asp Val His Ala Lys Asp Lys Gly Gly Leu Val Pro Leu His Asn
275 280 285

Ala Cys Ser Tyr Gly His Tyr Glu Val Ala Glu Leu Leu Val Lys His
290 295 300

Gly Ala Val Val Asn Val Ala Asp Leu Trp Lys Phe Thr Pro Leu His
305 310 315 320

Glu Ala Ala Ala Lys Gly Lys Tyr Glu Ile Cys Lys Leu Leu Leu Gln
325 330 335

His Gly Ala Asp Pro Thr Lys Lys Asn Arg Asp Gly Asn Thr Pro Leu
340 345 350

Asp Leu Val Lys Asp Gly Asp Thr Asp Ile
355 360

<210> 161
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 161
cgtcgaccca tggcggagtc ttcggataag ctctatcga 39

<210> 162
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 162
ggaaacgcgt ttggtgccag gatttactgt cagcttctt 39

<210> 163
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 163
cttaaacgcg ttgaaggaca aacaccttta gatttagtt 39

<210> 164
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 164
gtcgaaagcg gccgcttagc ctccgaactg tggatgcctc cacgctccat cgaccatacc 60
ttcaggcctc ataatctgg 79

<210> 165
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 165
tttgttcgcc cagactc 17

<210> 166
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 166
tatgtttcag gttcaggggg ag 22

<210> 167
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 167
gcggaagctg gaggagtgac 20

<210> 168
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 168
gtcactcctc cagcttccgc 20

<210> 169
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 169
aagccctgaa gaagcagctc 20

<210> 170
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 170

gagctgcttc ttcagggctt 20

<210> 171
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 171
cagacaccca accggaagga 20

<210> 172
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 172
tccttccggt tgggtgtctg 20

<210> 173
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 173
tccgcctcca ccaagagcct 20

<210> 174
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 174
aggctcttgg tggaggcgga 20

<210> 175
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 175
tggcctggtg gacatcgtta 20

<210> 176
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

```
<400> 176
taacgatgtc caccaggcca                                                    20


<210> 177
<211> 3308
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Parp1a-Tank2b
      Fusion

<220>
<221> CDS
<222> (1)..(3297)

<400> 177
atg aga ggc tcc cat cac cat cac cat cac gat tac gat atc cca acg      48
Met Arg Gly Ser His His His His His His Asp Tyr Asp Ile Pro Thr
  1               5                  10                  15

acc gaa aac ctg tat ttt cag ggc gcc atg gat ccg gaa ttc aaa ggc      96
Thr Glu Asn Leu Tyr Phe Gln Gly Ala Met Asp Pro Glu Phe Lys Gly
             20                  25                  30

cta cgt cga ccc atg gcg gag tct tcg gat aag ctc tat cga gtc gag     144
Leu Arg Arg Pro Met Ala Glu Ser Ser Asp Lys Leu Tyr Arg Val Glu
         35                  40                  45

tac gcc aag agc ggg cgc gcc tct tgc aag aaa tgt agc gag agc atc     192
Tyr Ala Lys Ser Gly Arg Ala Ser Cys Lys Lys Cys Ser Glu Ser Ile
     50                  55                  60

ccc aag gac tcg ctc cgg atg gcc atc atg gtg cag tcg ccc atg ttt     240
Pro Lys Asp Ser Leu Arg Met Ala Ile Met Val Gln Ser Pro Met Phe
 65                  70                  75                  80

gat gga aaa gtc cca cac tgg tac cac ttc tcc tgc ttc tgg aag gtg     288
Asp Gly Lys Val Pro His Trp Tyr His Phe Ser Cys Phe Trp Lys Val
                 85                  90                  95

ggc cac tcc atc cgg cac cct gac gtt gag gtg gat ggg ttc tct gag     336
Gly His Ser Ile Arg His Pro Asp Val Glu Val Asp Gly Phe Ser Glu
            100                 105                 110

ctt cgg tgg gat gac cag cag aaa gtc aag aag aca gcg gaa gct gga     384
Leu Arg Trp Asp Asp Gln Gln Lys Val Lys Lys Thr Ala Glu Ala Gly
        115                 120                 125

gga gtg aca ggc aaa ggc cag gat gga att ggt agc aag gca gag aag     432
Gly Val Thr Gly Lys Gly Gln Asp Gly Ile Gly Ser Lys Ala Glu Lys
    130                 135                 140

act ctg ggt gac ttt gca gca gag tat gtc aag tcc aac aga agt acg     480
Thr Leu Gly Asp Phe Ala Ala Glu Tyr Val Lys Ser Asn Arg Ser Thr
145                 150                 155                 160

tgc aag ggg tgt atg gag aag ata gaa aag ggc cag gtg cgc ctg tcc     528
Cys Lys Gly Cys Met Glu Lys Ile Glu Lys Gly Gln Val Arg Leu Ser
                165                 170                 175

aag aag atg gtg gac ccg gag aag cca cag cta ggc atg att gac cgc     576
Lys Lys Met Val Asp Pro Glu Lys Pro Gln Leu Gly Met Ile Asp Arg
            180                 185                 190

tgg tac cat cca ggc tgc ttt gtc aag aac agg gag gag ctg ggt ttc     624
Trp Tyr His Pro Gly Cys Phe Val Lys Asn Arg Glu Glu Leu Gly Phe
        195                 200                 205
```

```
cgg ccc gag tac agt gcg agt cag ctc aag ggc ttc agc ctc ctt gct     672
Arg Pro Glu Tyr Ser Ala Ser Gln Leu Lys Gly Phe Ser Leu Leu Ala
    210                 215                 220

aca gag gat aaa gaa gcc ctg aag aag cag ctc cca gga gtc aag agt     720
Thr Glu Asp Lys Glu Ala Leu Lys Lys Gln Leu Pro Gly Val Lys Ser
225                 230                 235                 240

gaa gga aag aga aaa ggc gat gag gtg gat gga gtg gat gaa gtg gcg     768
Glu Gly Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala
                245                 250                 255

aag aag aaa tct aaa aaa gaa aaa gac aag gat agt aag ctt gaa aaa     816
Lys Lys Lys Ser Lys Lys Glu Lys Asp Lys Asp Ser Lys Leu Glu Lys
            260                 265                 270

gcc cta aag gct cag aac gac ctg atc tgg aac atc aag gac gag cta     864
Ala Leu Lys Ala Gln Asn Asp Leu Ile Trp Asn Ile Lys Asp Glu Leu
        275                 280                 285

aag aaa gtg tgt tca act aat gac ctg aag gag cta ctc atc ttc aac     912
Lys Lys Val Cys Ser Thr Asn Asp Leu Lys Glu Leu Leu Ile Phe Asn
    290                 295                 300

aag cag caa gtg cct tct ggg gag tcg gcg atc ttg gac cga gta gct     960
Lys Gln Gln Val Pro Ser Gly Glu Ser Ala Ile Leu Asp Arg Val Ala
305                 310                 315                 320

gat ggc atg gtg ttc ggt gcc ctc ctt ccc tgc gag gaa tgc tcg ggt     1008
Asp Gly Met Val Phe Gly Ala Leu Leu Pro Cys Glu Glu Cys Ser Gly
                325                 330                 335

cag ctg gtc ttc aag agc gat gcc tat tac tgc act ggg gac gtc act     1056
Gln Leu Val Phe Lys Ser Asp Ala Tyr Tyr Cys Thr Gly Asp Val Thr
            340                 345                 350

gcc tgg acc aag tgt atg gtc aag aca cag aca ccc aac cgg aag gag     1104
Ala Trp Thr Lys Cys Met Val Lys Thr Gln Thr Pro Asn Arg Lys Glu
        355                 360                 365

tgg gta acc cca aag gaa ttc cga gaa atc tct tac ctc aag aaa ttg     1152
Trp Val Thr Pro Lys Glu Phe Arg Glu Ile Ser Tyr Leu Lys Lys Leu
    370                 375                 380

aag gtt aaa aag cag gac cgt ata ttc ccc cca gaa acc agc gcc tcc     1200
Lys Val Lys Lys Gln Asp Arg Ile Phe Pro Pro Glu Thr Ser Ala Ser
385                 390                 395                 400

gtg gcg gcc acg cct ccg ccc tcc aca gcc tcg gct cct gct gct gtg     1248
Val Ala Ala Thr Pro Pro Pro Ser Thr Ala Ser Ala Pro Ala Ala Val
                405                 410                 415

aac tcc tct gct tca gca gat aag cca tta tcc aac atg aag atc ctg     1296
Asn Ser Ser Ala Ser Ala Asp Lys Pro Leu Ser Asn Met Lys Ile Leu
            420                 425                 430

act ctc ggg aag ctg tcc cgg aac aag gat gaa gtg aag gcc atg att     1344
Thr Leu Gly Lys Leu Ser Arg Asn Lys Asp Glu Val Lys Ala Met Ile
        435                 440                 445

gag aaa ctc ggg ggg aag ttg acg ggg acg gcc aac aag gct tcc ctg     1392
Glu Lys Leu Gly Gly Lys Leu Thr Gly Thr Ala Asn Lys Ala Ser Leu
    450                 455                 460

tgc atc agc acc aaa aag gag gtg gaa aag atg aat aag aag atg gag     1440
Cys Ile Ser Thr Lys Lys Glu Val Glu Lys Met Asn Lys Lys Met Glu
465                 470                 475                 480
```

gaa gta aag gaa gcc aac atc cga gtt gtg tct gag gac ttc ctc cag 1488
Glu Val Lys Glu Ala Asn Ile Arg Val Val Ser Glu Asp Phe Leu Gln
485 490 495

gac gtc tcc gcc tcc acc aag agc ctt cag gag ttg ttc tta gcg cac 1536
Asp Val Ser Ala Ser Thr Lys Ser Leu Gln Glu Leu Phe Leu Ala His
500 505 510

atc ttg tcc cct tgg ggg gca gag gtg aag gca gag cct gtt gaa gtt 1584
Ile Leu Ser Pro Trp Gly Ala Glu Val Lys Ala Glu Pro Val Glu Val
515 520 525

gtg gcc cca aga ggg aag tca ggg gct gcg ctc tcc aaa aaa agc aag 1632
Val Ala Pro Arg Gly Lys Ser Gly Ala Ala Leu Ser Lys Lys Ser Lys
530 535 540

ggc cag gtc aag gag gaa ggt atc aac aaa tct gaa aag aga atg aaa 1680
Gly Gln Val Lys Glu Glu Gly Ile Asn Lys Ser Glu Lys Arg Met Lys
545 550 555 560

tta act ctt aaa gga gga gca gct gtg gat cct gat tct gga ctg gaa 1728
Leu Thr Leu Lys Gly Gly Ala Ala Val Asp Pro Asp Ser Gly Leu Glu
565 570 575

cac tct gcg cat gtc ctg gag aaa ggt ggg aag gtc ttc agt gcc acc 1776
His Ser Ala His Val Leu Glu Lys Gly Gly Lys Val Phe Ser Ala Thr
580 585 590

ctt ggc ctg gtg gac atc gtt aaa gga acc aac tcc tac tac aag ctg 1824
Leu Gly Leu Val Asp Ile Val Lys Gly Thr Asn Ser Tyr Tyr Lys Leu
595 600 605

cag ctt ctg gag gac gac aag gaa aac agg tat tgg ata ttc agg tcc 1872
Gln Leu Leu Glu Asp Asp Lys Glu Asn Arg Tyr Trp Ile Phe Arg Ser
610 615 620

tgg ggc cgt gtg ggt acg gtg atc ggt agc aac aaa ctg gaa cag atg 1920
Trp Gly Arg Val Gly Thr Val Ile Gly Ser Asn Lys Leu Glu Gln Met
625 630 635 640

ccg tcc aag gag gat gcc att gag cac ttc atg aaa tta tat gaa gaa 1968
Pro Ser Lys Glu Asp Ala Ile Glu His Phe Met Lys Leu Tyr Glu Glu
645 650 655

aaa acc ggg aac gct tgg cac tcc aaa aat ttc acg aag tat ccc aaa 2016
Lys Thr Gly Asn Ala Trp His Ser Lys Asn Phe Thr Lys Tyr Pro Lys
660 665 670

aag ttc tac ccc ctg gag att gac tat ggc cag gat gaa gag gca gtg 2064
Lys Phe Tyr Pro Leu Glu Ile Asp Tyr Gly Gln Asp Glu Glu Ala Val
675 680 685

aag aag ctg aca gta aat cct ggc acc aaa cgc gtt gaa gga caa aca 2112
Lys Lys Leu Thr Val Asn Pro Gly Thr Lys Arg Val Glu Gly Gln Thr
690 695 700

cct tta gat tta gtt tca gca gat gat gtc agc gct ctt ctg aca gca 2160
Pro Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala
705 710 715 720

gcc atg ccc cca tct gct ctg ccc tct tgt tac aag cct caa gtg ctc 2208
Ala Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu
725 730 735

aat ggt gtg aga agc cca gga gcc act gca gat gct ctc tct tca ggt 2256
Asn Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly
740 745 750

```
cca tct agc cca tca agc ctt tct gca gcc agc agt ctt gac aac tta    2304
Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu
        755                 760                 765

tct ggg agt ttt tca gaa ctg tct tca gta gtt agt tca agt gga aca    2352
Ser Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr
    770                 775                 780

gag ggt gct tcc agt ttg gag aaa aag gag gtt cca gga gta gat ttt    2400
Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe
785                 790                 795                 800

agc ata act caa ttc gta agg aat ctt gga ctt gag cac cta atg gat    2448
Ser Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp
                805                 810                 815

ata ttt gag aga gaa cag atc act ttg gat gta tta gtt gag atg ggg    2496
Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly
            820                 825                 830

cac aag gag ctg aag gag att gga atc aat gct tat gga cat agg cac    2544
His Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His
        835                 840                 845

aaa cta att aaa gga gtc gag aga ctt atc tcc gga caa caa ggt ctt    2592
Lys Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu
    850                 855                 860

aac cca tat tta act ttg aac acc tct ggt agt gga aca att ctt ata    2640
Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile
865                 870                 875                 880

gat ctg tct cct gat gat aaa gag ttt cag tct gtg gag gaa gag atg    2688
Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met
                885                 890                 895

caa agt aca gtt cga gag cac aga gat gga ggt cat gca ggt gga atc    2736
Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile
            900                 905                 910

ttc aac aga tac aat att ctc aag att cag aag gtt tgt aac aag aaa    2784
Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys
        915                 920                 925

cta tgg gaa aga tac act cac cgg aga aaa gaa gtt tct gaa gaa aac    2832
Leu Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn
    930                 935                 940

cac aac cat gcc aat gaa cga atg cta ttt cat ggg tct cct ttt gtg    2880
His Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val
945                 950                 955                 960

aat gca att atc cac aaa ggc ttt gat gaa agg cat gcg tac ata ggt    2928
Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly
                965                 970                 975

ggt atg ttt gga gct ggc att tat ttt gct gaa aac tct tcc aaa agc    2976
Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser
            980                 985                 990

aat caa tat gta tat gga att gga gga ggt act ggg tgt cca gtt cac    3024
Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His
        995                1000                1005

aaa gac aga tct tgt tac att tgc cac agg cag ctg ctc ttt tgc cgg    3072
Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg
   1010                1015                1020
```

gta acc ttg gga aag tct ttc ctg cag ttc agt gca atg aaa atg gca 3120
Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala
1025 1030 1035 1040

cat tct cct cca ggt cat cac tca gtc act ggt agg ccc agt gta aat 3168
His Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn
1045 1050 1055

ggc cta gca tta gct gaa tat gtt att tac aga gga gaa cag gct tat 3216
Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr
1060 1065 1070

cct gag tat tta att act tac cag att atg agg cct gaa ggt atg gtc 3264
Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val
1075 1080 1085

gat gga gcg tgg agg cat cca cag ttc gga ggc taagcggccg c 3308
Asp Gly Ala Trp Arg His Pro Gln Phe Gly Gly
1090 1095

<210> 178
<211> 1099
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence:Parp1a-Tank2b
Fusion

<400> 178
Met Arg Gly Ser His His His His His His Asp Tyr Asp Ile Pro Thr
1 5 10 15

Thr Glu Asn Leu Tyr Phe Gln Gly Ala Met Asp Pro Glu Phe Lys Gly
20 25 30

Leu Arg Arg Pro Met Ala Glu Ser Ser Asp Lys Leu Tyr Arg Val Glu
35 40 45

Tyr Ala Lys Ser Gly Arg Ala Ser Cys Lys Lys Cys Ser Glu Ser Ile
50 55 60

Pro Lys Asp Ser Leu Arg Met Ala Ile Met Val Gln Ser Pro Met Phe
65 70 75 80

Asp Gly Lys Val Pro His Trp Tyr His Phe Ser Cys Phe Trp Lys Val
85 90 95

Gly His Ser Ile Arg His Pro Asp Val Glu Val Asp Gly Phe Ser Glu
100 105 110

Leu Arg Trp Asp Asp Gln Gln Lys Val Lys Lys Thr Ala Glu Ala Gly
115 120 125

Gly Val Thr Gly Lys Gly Gln Asp Gly Ile Gly Ser Lys Ala Glu Lys
130 135 140

Thr Leu Gly Asp Phe Ala Ala Glu Tyr Val Lys Ser Asn Arg Ser Thr
145 150 155 160

Cys Lys Gly Cys Met Glu Lys Ile Glu Lys Gly Gln Val Arg Leu Ser
165 170 175

Lys Lys Met Val Asp Pro Glu Lys Pro Gln Leu Gly Met Ile Asp Arg
180 185 190

Trp Tyr His Pro Gly Cys Phe Val Lys Asn Arg Glu Glu Leu Gly Phe
195 200 205

Arg Pro Glu Tyr Ser Ala Ser Gln Leu Lys Gly Phe Ser Leu Leu Ala
210 215 220

Thr Glu Asp Lys Glu Ala Leu Lys Lys Gln Leu Pro Gly Val Lys Ser
225 230 235 240

Glu Gly Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala
245 250 255

Lys Lys Lys Ser Lys Lys Glu Lys Asp Lys Asp Ser Lys Leu Glu Lys
260 265 270

Ala Leu Lys Ala Gln Asn Asp Leu Ile Trp Asn Ile Lys Asp Glu Leu
275 280 285

Lys Lys Val Cys Ser Thr Asn Asp Leu Lys Glu Leu Leu Ile Phe Asn
290 295 300

Lys Gln Gln Val Pro Ser Gly Glu Ser Ala Ile Leu Asp Arg Val Ala
305 310 315 320

Asp Gly Met Val Phe Gly Ala Leu Leu Pro Cys Glu Glu Cys Ser Gly
325 330 335

Gln Leu Val Phe Lys Ser Asp Ala Tyr Tyr Cys Thr Gly Asp Val Thr
340 345 350

Ala Trp Thr Lys Cys Met Val Lys Thr Gln Thr Pro Asn Arg Lys Glu
355 360 365

Trp Val Thr Pro Lys Glu Phe Arg Glu Ile Ser Tyr Leu Lys Lys Leu
370 375 380

Lys Val Lys Lys Gln Asp Arg Ile Phe Pro Pro Glu Thr Ser Ala Ser
385 390 395 400

Val Ala Ala Thr Pro Pro Pro Ser Thr Ala Ser Ala Pro Ala Ala Val
405 410 415

Asn Ser Ser Ala Ser Ala Asp Lys Pro Leu Ser Asn Met Lys Ile Leu
420 425 430

Thr Leu Gly Lys Leu Ser Arg Asn Lys Asp Glu Val Lys Ala Met Ile
435 440 445

Glu Lys Leu Gly Gly Lys Leu Thr Gly Thr Ala Asn Lys Ala Ser Leu
450 455 460

Cys Ile Ser Thr Lys Lys Glu Val Glu Lys Met Asn Lys Lys Met Glu
465 470 475 480

Glu Val Lys Glu Ala Asn Ile Arg Val Val Ser Glu Asp Phe Leu Gln
485 490 495

Asp Val Ser Ala Ser Thr Lys Ser Leu Gln Glu Leu Phe Leu Ala His
500 505 510

Ile Leu Ser Pro Trp Gly Ala Glu Val Lys Ala Glu Pro Val Glu Val
515 520 525

Val Ala Pro Arg Gly Lys Ser Gly Ala Ala Leu Ser Lys Lys Ser Lys
530 535 540

Gly Gln Val Lys Glu Glu Gly Ile Asn Lys Ser Glu Lys Arg Met Lys
545 550 555 560

Leu Thr Leu Lys Gly Gly Ala Ala Val Asp Pro Asp Ser Gly Leu Glu
565 570 575

His Ser Ala His Val Leu Glu Lys Gly Gly Lys Val Phe Ser Ala Thr
580 585 590

```
Leu Gly Leu Val Asp Ile Val Lys Gly Thr Asn Ser Tyr Tyr Lys Leu
        595                 600                 605

Gln Leu Leu Glu Asp Asp Lys Glu Asn Arg Tyr Trp Ile Phe Arg Ser
    610                 615                 620

Trp Gly Arg Val Gly Thr Val Ile Gly Ser Asn Lys Leu Glu Gln Met
625                 630                 635                 640

Pro Ser Lys Glu Asp Ala Ile Glu His Phe Met Lys Leu Tyr Glu Glu
                645                 650                 655

Lys Thr Gly Asn Ala Trp His Ser Lys Asn Phe Thr Lys Tyr Pro Lys
            660                 665                 670

Lys Phe Tyr Pro Leu Glu Ile Asp Tyr Gly Gln Asp Glu Glu Ala Val
        675                 680                 685

Lys Lys Leu Thr Val Asn Pro Gly Thr Lys Arg Val Glu Gly Gln Thr
    690                 695                 700

Pro Leu Asp Leu Val Ser Ala Asp Asp Val Ser Ala Leu Leu Thr Ala
705                 710                 715                 720

Ala Met Pro Pro Ser Ala Leu Pro Ser Cys Tyr Lys Pro Gln Val Leu
                725                 730                 735

Asn Gly Val Arg Ser Pro Gly Ala Thr Ala Asp Ala Leu Ser Ser Gly
            740                 745                 750

Pro Ser Ser Pro Ser Ser Leu Ser Ala Ala Ser Ser Leu Asp Asn Leu
        755                 760                 765

Ser Gly Ser Phe Ser Glu Leu Ser Ser Val Val Ser Ser Ser Gly Thr
    770                 775                 780

Glu Gly Ala Ser Ser Leu Glu Lys Lys Glu Val Pro Gly Val Asp Phe
785                 790                 795                 800

Ser Ile Thr Gln Phe Val Arg Asn Leu Gly Leu Glu His Leu Met Asp
                805                 810                 815

Ile Phe Glu Arg Glu Gln Ile Thr Leu Asp Val Leu Val Glu Met Gly
            820                 825                 830

His Lys Glu Leu Lys Glu Ile Gly Ile Asn Ala Tyr Gly His Arg His
        835                 840                 845

Lys Leu Ile Lys Gly Val Glu Arg Leu Ile Ser Gly Gln Gln Gly Leu
    850                 855                 860

Asn Pro Tyr Leu Thr Leu Asn Thr Ser Gly Ser Gly Thr Ile Leu Ile
865                 870                 875                 880

Asp Leu Ser Pro Asp Asp Lys Glu Phe Gln Ser Val Glu Glu Glu Met
                885                 890                 895

Gln Ser Thr Val Arg Glu His Arg Asp Gly Gly His Ala Gly Gly Ile
            900                 905                 910

Phe Asn Arg Tyr Asn Ile Leu Lys Ile Gln Lys Val Cys Asn Lys Lys
        915                 920                 925

Leu Trp Glu Arg Tyr Thr His Arg Arg Lys Glu Val Ser Glu Glu Asn
    930                 935                 940

His Asn His Ala Asn Glu Arg Met Leu Phe His Gly Ser Pro Phe Val
945                 950                 955                 960
```

Asn Ala Ile Ile His Lys Gly Phe Asp Glu Arg His Ala Tyr Ile Gly
965 970 975

Gly Met Phe Gly Ala Gly Ile Tyr Phe Ala Glu Asn Ser Ser Lys Ser
980 985 990

Asn Gln Tyr Val Tyr Gly Ile Gly Gly Gly Thr Gly Cys Pro Val His
995 1000 1005

Lys Asp Arg Ser Cys Tyr Ile Cys His Arg Gln Leu Leu Phe Cys Arg
1010 1015 1020

Val Thr Leu Gly Lys Ser Phe Leu Gln Phe Ser Ala Met Lys Met Ala
1025 1030 1035 1040

His Ser Pro Pro Gly His His Ser Val Thr Gly Arg Pro Ser Val Asn
1045 1050 1055

Gly Leu Ala Leu Ala Glu Tyr Val Ile Tyr Arg Gly Glu Gln Ala Tyr
1060 1065 1070

Pro Glu Tyr Leu Ile Thr Tyr Gln Ile Met Arg Pro Glu Gly Met Val
1075 1080 1085

Asp Gly Ala Trp Arg His Pro Gln Phe Gly Gly
1090 1095

INTERNATIONAL SEARCH REPORT

Int[redacted]tional Application No
PCT/US 00/17827

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/54 C12N9/10 C07K16/40 C12Q1/68 C12Q1/48
A61K38/45 //A61P35/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 15647 A (GARVAN INST MED RES ;SUTHERLAND ROBERT LYNDSAY (AU); DALY ROGER JO) 1 April 1999 (1999-04-01) page 8 -page 10, line 24 page 13 -page 17 ---	7
A	SMITH S ET AL: "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres" SCIENCE, vol. 282, no. 5393, 20 November 1998 (1998-11-20), pages 1484-1487, XP002118903 ISSN: 0036-8075 cited in the application ---	
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 December 2000	11/12/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Andres, S

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 00 61813 A (FUNK WALTER D ;MORIN GREGG B (US); GERON CORP (US); PIATYSZEK MIEC) 19 October 2000 (2000-10-19) page 2, line 9 -page 3, line 15 examples claims figure 4 -----	1,3,4, 6-17, 19-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

In[redacted]tional Application No
PCT/US 00/17827

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9915647 A	01-04-1999	AU 9245898 A	12-04-1999
		EP 1017802 A	12-07-2000
WO 0061813 A	19-10-2000	NONE	

Form PCT/ISA/210 (patent family annex) (July 1992)

This Page Blank (usp

This Page Blank (uspto)

This Page Blank (usp